DOCUMENT RESUME

ED 459 523 EA 031 473

AUTHOR Jandris, Thomas P.
TITLE Data-Based Decision-Making. Essentials for Principals[TM].
INSTITUTION Educational Research Service, Arlington, VA.; National Association of Elementary School Principals, Alexandria, VA.
PUB DATE 2001-00-00
NOTE 102p.
AVAILABLE FROM National Association of Elementary School Principals, 1615 Duke St., Alexandria, VA 22314-3483 (Order #EDBDM, $24.95). Tel: 800-386-2377 (Toll Free); Fax: 800-396-2377 (Toll Free); Web site: http://www.naesp.org.
PUB TYPE Guides - Non-Classroom (055)
EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Academic Achievement; Administration; Administrator Guides; Data Analysis; *Decision Making; *Educational Assessment; Elementary Secondary Education; *Evaluation; *Instructional Improvement; Intervention; Program Implementation

ABSTRACT

The standards and accountability movement of the last decade has resulted in the collection of more assessment data about students than ever before. Often, the results from high-stakes tests have been used merely to grade and rank schools and students. A new trend, however, has emerged where school leaders see the power of using these same assessment data to improve instruction at the school, classroom, and student levels. This publication is designed to give school leaders a detailed understanding of assessment so they and their schools can make the best possible use of this tool for educational improvement. Information is included about the "big picture" of assessment today; implementation of an effective schoolwide assessment program; the assessment "process," including aspects such as reflecting on the data and providing interventions based on assessment results; ways to enhance the effectiveness of classroom assessments in providing useful data; the importance of incorporating data into a cycle of continuous improvement; and the use of technology to strengthen the assessment process. The publication ends with a list of Web resources for using assessment data, a list of recommended books and articles, and an assessment glossary. (Contains 99 references.) (RT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 459 523

ssentials

Data-Based Decision-Making

Principals

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

W. T. Greenleaf

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

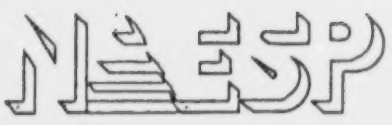
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

EA 031473

Essentials for Principals™

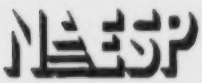
Data-Based Decision-Making



Essentials for Principals: Data-Based Decision-Making was prepared by Educational Research Service for the National Association of Elementary School Principals.
Copyright © 2001 National Association of Elementary School Principals.

The *Essentials for Principals* series is produced under the direction of Ernie Mannino, Assistant Executive Director, National Principals Resource Center. Mr. Mannino has guided the development of the series, providing direction and offering suggestions to ensure that the publications are quality resources for principals.

For information about how to order additional copies of this publication, contact NAESP at the address below.



National Association of Elementary School Principals

1615 Duke Street, Alexandria, VA 22314-3483. Phone: (703) 684-3345. Fax: (703) 549-5568. Email: naesp@naesp.org. Website: www.naesp.org.

The 28,500 members of the National Association of Elementary School Principals provide administrative and instructional leadership for public and private elementary and middle schools throughout the United States, Canada, and overseas. Founded in 1921, NAESP is today a vigorously independent professional association with its own headquarters building in Alexandria, Virginia, just across the Potomac River from the nation's capital. From this special vantage point, NAESP conveys the unique perspective of the elementary and middle school principal to the highest policy councils of our national government. Through national and regional meetings, award-winning publications, and joint efforts with its 50 state affiliates, NAESP is a strong advocate both for its members and for the 33 million American children enrolled in preschool, kindergarten, and grades 1 through 8.

Educational Research Service

2000 Clarendon Boulevard, Arlington, VA 22201-2908. Phone: (703) 243-2100 or (800) 791-9308. Fax: (703) 243-5971 or (800) 791-9309. Email: ers@ers.org. Website: www.ers.org.

Educational Research Service is the nonprofit foundation serving the research and information needs of the nation's education leaders and the public. ERS was founded jointly by seven national school management associations, including NAESP. ERS serves as both a national source and a clearinghouse for school research and information. By compiling, analyzing, and sharing information that is essential to effective decision making, ERS assists local school district administrators with both day-to-day operations and long-range planning. It provides a unique approach to gathering data and reporting information through services that are geared to a quick, yet thorough, response to subscriber requests.

Author: Thomas P. Jandris

Thomas P. Jandris is Director of State Services for the Education Commission of the States (ECS), where he and his team are the consultants and technical advisors to our nation's governors, legislators, and other education leaders. In his role as Director, Dr. Jandris is the chief architect and leader of the Urban Coalition and Corporate Partners Programs at ECS.

Prior to his tenure at ECS, Dr. Jandris had extensive careers in both education and the private sector. He was a teacher, principal, and director of secondary education at a large urban district. He founded and grew a highly successful human resources consulting company. He is a past U.S. Olympic Team member and syndicated columnist.

Dr. Jandris is a licensed psychologist who continues to teach at the graduate level. His education includes a Ph.D. from the University of Minnesota, an M.Ed. from Wayne State, and a B.S. in Education from Eastern Illinois University.

Contents

Foreword v

About this *Essentials for Principals* Publication 1

Assessment: The Big Picture 3
- What Is Assessment? 3
  - What Does Assessment Look Like? 3
- Purposes of Assessment 4
  - Example: Classroom Use of Multiple Assessment Methods 5
- The New Era of Standards and Assessment 6
  - Charting the Purposes of Assessment 7
  - Tool for Curriculum Alignment 11
  - Example: Basing Assessment on Learning Objectives 12
  - Classroom Instruction with Assessment in the Continuous Cycle 13
  - Example: Using Assessment Data to Make Instructional Decisions 14
- Why Put Time and Effort into Making Effective Use of Assessment Data? 15
- Making It Work: Expectations, Instruction, and Assessment 16
  - Alignment Model 16

Creating Schoolwide Conditions for Effective Use of Assessment Data 18
  - The Strategies 19
- Strategy 1: Supportive School Climate 19
  - The Inquiry Cycle: Continuous, Reflective Use of Data 20
- Strategy 2: Time for Assessment 21
  - Scheduling Approaches for Teacher Collaboration in Four Schools 22
  - Example: Teacher Collaboration to Develop Assessments 23
- Strategy 3: Teacher Training 24
- Strategy 4: Stating Expectations for Performance 24
  - Assessment Literacy Self-Appraisal 25
  - Some Links to State Standards 26
  - Example: Communicating with the Community about the Purposes of Assessment 27

The Process of Assessment 29
- Collecting the Data: Summative and Formative Assessment 29

Indicates a special feature.

More Information on Formative Assessment 30
Looking at the Data and Understanding It 31
Questions to Ask 32
Example: Partial SAT-9 Analysis for One Class 34
Example: Using Modeling and Frequent Feedback to Improve Performance Levels 35
Guidelines for Looking at Student Work 37
Intervention: Taking Action Based on the Data 38
Using Assessment Data to Identify Good Teaching 40
Resources for Accountability 40

Enhancing Classroom Assessment 42
Authentic Assessment 42
More Information on Authentic Assessment 43
Multiple Measures of Performance 43
Example: Using Multiple Measures of Performance 44
What About the Standardized Test My District Gives Every Year? 45
Aligning Assessment Methods with Purposes 46
Standardized Tests vs. Diagnostic Assessments 47
Assessment Toolbox 47

The Continuous Improvement Process 49
Communicating Achievement and Performance Targets 50
Using Data to Determine Whether Standards Have Been Met 51
Guiding Assumptions of Data-Based Decision-Making 53
Basic Data Elements 54
The Continuous Improvement Cycle 55
The Brazosport 8-Step Instructional Process 56

The Promise of Technology to Improve Use of Assessment Data 61
How Technology-Savvy Are Your Teachers' Portfolios? 62

Conclusion 64

Web Resources for Using Assessment Data 66

Recommended Books and Articles 69

An Assessment Glossary 71

References 88

Foreword

As the author of this publication points out in his introduction, schools and their leaders today are collecting more assessment data on their students than they have ever done before. New state-level mandates have been added to many districts' already-full assessment schedules. The day-to-day assessments that teachers carry out as an essential part of the instructional process are just one part of the larger picture.

Discovering how to use the wealth of data on student achievement in positive and productive ways can be an important advantage for the effective principal. This *Essentials for Principals* publication, *Data-Based Decision-Making,* presents research and practical experience in a down-to-earth discussion of what every school leader should know about the topic.

Each publication in the *Essentials for Principals* series is jointly developed by the National Association of Elementary School Principals and Educational Research Service to provide information that is both practical and solidly research-based on issues of vital importance to school leaders. Like the previous volumes, this new publication in the series is a valuable resource for both beginning and experienced principals. In addition to the many examples and lists of additional resources included in the publication, this *Essentials* guide includes an extensive glossary of terms related to assessment and assessment data.

NAESP and ERS believe that this *Essentials* guide will help principals provide the best leadership possible to help their staffs use assessment data as a powerful tool to improve instruction and promote high levels of student learning. We welcome your comments and suggestions about this publication, as well as ways in which the *Essentials for Principals* series can be most helpful to you.

Vincent L. Ferrandino
Executive Director
National Association of
Elementary School Principals

John M. Forsyth
President
Educational Research Service

About this *Essentials for Principals* Publication

The standards and accountability movement of the last decade has resulted in the collection of more assessment data about students than ever before. Often, the results from high-stakes tests have been used merely to grade and rank schools and students. Used solely this way, assessment data is of doubtful value in improving student achievement. Educators know this, and many of them have therefore viewed the imposition of state-mandated tests as an unwanted burden that actually stands in the way of school improvement.

But a new trend is now starting to emerge. Education leaders are beginning to see the power of using this same assessment data, along with data gathered from carefully designed classroom-based assessments, to improve instruction at the school, classroom, and student levels. This *Essentials for Principals* publication is designed to give you, as the principal, the detailed understanding of assessment you need to ensure that your school makes the best possible use of this valuable tool for educational improvement.

Although the standardized tests now mandated by states across the nation receive the most attention in the media and among many education policy makers today, assessment spans a whole range of activities and has always been an integral part of curriculum and instruction in public schools. Assessment data comes from many sources. That's why this publication does not focus on standardized assessments, but rather gives an overview of the use of all assessment data—and especially classroom-based assessment data—to improve instruction. Included is information about:

- the "big picture" of assessment today;
- implementation of an effective schoolwide assessment program;
- the assessment "process," including aspects such as reflecting on the data and providing interventions based on assessment results;

- ways to enhance the effectiveness of classroom assessments in providing useful data;
- the importance of incorporating data into a cycle of continuous improvement; and
- the use of technology to strengthen the assessment process.

If assessment is "embedded in instruction" as this *Essentials* publication recommends, and if the data is carefully analyzed by the principal, teachers, and other school staff, the process can be a powerful force for improving instruction and raising student achievement.

Assessment: The Big Picture

What Is Assessment?

Assessment is a tool. Teachers use this tool to collect data to better understand the current knowledge and skills that individual students possess. Principals and other school leaders use it similarly, but they can also use it more broadly—for example, to evaluate the effectiveness of instruction or to improve the performance of cohorts of students. When analyzed and interpreted, the information generated by an assessment can be used to make crucial instructional decisions about pacing, remediation, and the introduction of new information, as well as staff development and program design.

There are a wide variety of assessment methods available for measuring student progress (see the list at the right). They can range from the teacher's subjective judgment of a student's participation in classroom discussions to the administration of a four-day standardized test.

At the classroom level, teachers can construct tests or use test items banks to assess student learning. The scope of a test can range from a few pages or single lesson, to a chapter or unit, to a whole semester's worth of material. Tests, however, are only one type of assessment.

What Does Assessment Look Like?

- Structured observations
- Spur-of-the-moment observations
- Checklists
- Rating scales
- Formal interviews
- Conferences
- Portfolio collections of work samples
- Extended projects, products, and exhibitions
- Performance exams
- Constructed responses
- Essay questions
- Peer review
- Reflective writing
- Pop quizzes
- Teacher-made tests
- Tests accompanying the textbook
- Standardized tests
- Classroom questions
- Journaling
- Daily seatwork and homework assignments

Assessment may look like any other classroom activity.

Teachers can use rubrics to assess presentations, essays, and projects. Observations are also useful for assessing work skills and student interactions. All assessments, of course, should be linked to the objectives and standards of the instruction.

The educational literature often uses the terms *evaluation* and *assessment* as synonyms. There are differences between the two terms, although there is also a great deal of overlap.

Evaluation is a judgment of the value or merit of something (Scriven 1967). Evaluation is often applied to a large context such as a program (i.e., new math, whole language), but it is also applicable to an instructional method such as inquiry or cooperative learning.

Assessment has a much narrower focus and is almost always related to student achievement. Ormrod (2000) defines assessment as the collection of data about a student (i.e., tests, observations, projects, attitudes, and work habits) in order to draw inferences about the student's knowledge and abilities. Assessment, the collection of data, is often considered a part of an educational evaluation plan.

Assessment in the classroom typically focuses on student achievement—did a student or class master the objectives, or how well did they perform on the test? Evaluation asks broader questions that focus our attention on judging the worth of a program or instructional method. An evaluation might ask whether we are doing what we said we would do, whether students are learning what we stated they would, or whether a method is effective in improving student performance. Our focus in this book is on assessment and using assessment data to improve instruction.

Purposes of Assessment

Assessments are used for many purposes. Students, teachers, administrators, and parents have different perspectives and uses for assessment results.

Educators use assessment results to help improve teaching and learning and to evaluate instructional practices, programs, and schools. Assessments are also used to measure student achievement and evaluate individual students' mastery of skills so that students, parents, administrators, and teachers know how students are progressing toward agreed-upon standards. Assessments are used to rank students, classrooms, and schools, to generate comparative information, and to create accountability data. They are a major force in shaping public perception about the quality of our schools. Assessments generate the data on which instructional and policy decisions are made.

The heart of assessment is a continuing process in which the teacher, in collaboration with the student, uses information to guide the next steps in learning. The

Example: Classroom Use of Multiple Assessment Methods

Mr. Jones is teaching a unit on the short story. Yesterday, he gave a quiz asking students to define the components of a short story. Later in the unit, each student will write an essay on a specific author. The final project for the unit is team presentations on the students' analysis of several short stories.

The quizzes are easy to grade because they are a mix of short-answer and multiple-choice items. To grade the essays, Mr. Jones uses the district's rubric, which includes criteria for content, organization, style, and grammar. He has also added three additional items specific to the concept of the short story that he wants students to address. Mr. Jones has a set of similar rubrics for the team presentations, which includes criteria for the content, a PowerPoint presentation, and presentation skills.

As the students are working on the essays and presentations, Mr. Jones observes individual work habits and interpersonal skills. He makes anecdotal notes in a notebook. When he prepares the grades for the reporting period, he refers to the notebook to help him develop a rating for study and interpersonal skills. He also uses his notes to write comments to the students and parents on the report card.

This scenario from Mr. Jones' classroom provides an example of the different types of assessment we might find in a classroom. Most teachers base their grades and decisions on many sources of data rather than relying solely on test scores.

following are four ways of viewing the purposes of assessment, developed by Stiggins (1997).

- **Information for students, teachers, principals, and parents to inform classroom-evel decisions.** For the student, assessment provides feedback on understanding. Results of a test or comments on a written assignment will help the student answer such questions as "Am I getting it?" and "How am I doing?" Assessments can also help teachers answer related questions such as "Which students are learning?" Similarly, an analysis of the results can help a teacher pinpoint specific difficulties or misconceptions to further explain or correct. Teachers can also use assessment data for diagnostic purposes. When applied to the class, the data can answer such questions as, "Is the

class ready for a new unit?" A teacher who is using individualized instruction can use assessment data to determine where a student should start, and when he or she is ready for another unit.

- **The insight and motivation students (or teachers) need to aspire to academic excellence.** The knowledge of a correct or incorrect response is useful feedback for the student and the teacher. Feedback can help the student clarify understanding and identify misconceptions. Using this specific feedback, students can be motivated to focus on the relevant content either individually or in small groups to refine their understanding. Teachers can also use feedback to help the students understand expectations and develop a model of exemplary performance.
- **Evidence of student achievement for public accountability.** Assessments provide two levels of accountability. Day-to-day classroom assessments provide a microscopic view of the progress of individual students and the whole class. This kind of assessment can measure student progress from day to day, weekly, or over a grading period. More formal achievement tests typically provide a broader view of student achievement relative to standards. The timeframe for this view may vary from an annual assessment to one given every two or three years. Both levels of assessment, however, provide students, teachers, administrators, and parents with snapshots of individual student and class progress.
- **Information for internal programmatic decision-making.** When placed in the broader contexts of the course, grade, school, or district, assessment information is useful for identifying the strengths and weaknesses of the program. Thoughtful analysis and interpretation of this information can lead to improvements in the classroom and school environments that further enhance student learning (Assessment Training Institute 2001).

Given the different uses for assessment, it is critical that educators select the appropriate type of assessment to deliver useful, accurate, and reliable information. The chart on pages 8-10 illustrates these differences. This chart is helpful for planning assessment tools that will yield the appropriate data to address the questions education stakeholders need to answer.

The New Era of Standards and Assessment

During the latter half of the 20th century, the role of testing has changed. Two general areas of change are discussed here:

- **Assessments Linked to Curriculum Standards.** The kind of tests being emphasized has shifted from generic achievement tests, which were

designed to measure student knowledge and skills on a general body of knowledge, to assessments linked to local, state, and national curriculum standards.

- **Assessment Data as a Tool for Instructional Improvement.** At the same time, the purposes of assessment have come into better focus. Teacher-made tests are now expected to do more than simply provide a score for determining a grade—they must also be used to guide instruction for both the whole class and the individual student. State-mandated or district-designed tests must play a more positive role than that of assigning ranks and ratings to schools and educators—they must also be used to improve the instructional program.

Assessments Linked to Curriculum Standards

Before we can design, adapt, or use an assessment, we need to answer a key question: What are we measuring? The emphasis on common standards that should be taught and assessed is one of the major trends currently shaping assessment.

Standards are conceptually nothing new—but they have received a new emphasis over the last decade, through various state initiatives and through the passage of the Goals 2000: Educate America Act (see www.ed.gov/legislation/GOALS2000/TheAct/index.html).

During the last decade, nearly every state has been engaged in setting rigorous standards that define what students should know and be able to do for future success, and in developing assessment systems aligned to these high standards (Birenbaum and Douchy 1996). Additionally, standards have been set at the local school district level, nationally, or even internationally (Achieve, Inc. 1999).

However, the use of the term *standards* in the discussion of student evaluation can be confusing. As educators we talk of "raising standards," "developing a national standard," and "establishing standards to guide students and teachers." As these phrases suggest, the term *standards* can be used in more than one way. For example:

- When the term "standards" refers to consistent high expectations, it usually means that students are required to achieve to a certain level in a particular project, course, or grade. Standards specify, for example, what students should know and be able to do in core subjects at critical points in their educational career. They establish a rigorous level of knowledge and skills.
- The term "standards" can also refer to an exemplary performance in an area of study that will serve as a benchmark for students and teachers. Standards provide a model and criteria against which students and teachers can measure achievement during the learning process. When used in this

Charting the Purposes of Assessment				
Users and Uses of Standardized Assessment Results				
	Users	**Key Question(s) to Be Answered**	**Information Needed**	**Benefits**
Classroom/Instructional Level • Day-to-Day Classroom Assessments	Student	Am I meeting the teacher's standards? What help do I need to succeed? Are the results worth my investment of energy?	Continuous information about individual student attainment of specific instructional requirements	Self-evaluation and motivation Focus on content/ achievement Comparison to exemplary work
	Teacher	Which students need what help? Who among my students should work together? What grade should appear on the report card? Did my teaching strategies work? How can I become a better teacher?	Continuous information about individual student achievement Continuous assessment of group performance	Focus on students and learning Targeted reflection Diagnostic planning Individualized student adaptations and/or interventions Increased accountability data
	Parent	Is my child succeeding in school? What does my child need to succeed? Is my child's teacher(s) doing a good job? Is this district doing a good job?	Continuous feedback on the student's mastery of required material	Home/school communication Timely interventions possible Increased home-school alignment of perceived student achievement Accurate, understandable, usable information about achievement

Charting the Purposes of Assessment (continued)

Users and Uses of Standardized Assessment Results

INSTRUCTIONAL SUPPORT LEVEL • STANDARDIZED ASSESSMENTS

Users	Key Question(s) to Be Answered	Information Needed	Benefits
Principal/ Vice Principal	Is instruction in particular areas producing results? Is this teacher effective? What kinds of professional development will help? How shall we spend building resources to be effective?	Periodic assessment of group achievement	Focus on students, teachers, and learning Increased accountability data Targeted staff development Teacher interventions Program and school evaluation and planning
Mentor/ Lead	What does this teacher need to do the job?	Periodic assessment of group achievement	Focus on students, teachers, and learning Targeted peer coaching/ mentoring Timely interventions
Counselor	Who needs (can have access to) special support services? What students should be assigned to which teachers to optimize results?	Periodic assessment of individual achievement	Diagnostic information Support service interventions Focus on students
Curriculum Director	Is our program of instruction effective?	Periodic assessment of group achievement	Program evaluation/planning Accountability data Targeted staff development

Charting the Purposes of Assessment (continued)

Users and Uses of Standardized Assessment Results

Policy Level • Standardized Assessments

Users	Key Question(s) to Be Answered	Information Needed	Benefits
Superintendent	Are programs producing student learning? Is the building principal producing results? Which programs need/deserve more resources?	Periodic assessment of group mastery of district curriculum	Program/people interventions Program/system evaluation Accountability data
School Board	Are the students in the district learning? Is the superintendent producing results?	Periodic assessment of group achievement	Administrative interventions Program/system evaluation Accountability data
State Dept. of Ed	Are programs across the state producing results?	Periodic assessment of group mastery of state curriculum	Program/system evaluation Accountability/ needs data
Citizen/ Legislator (State/ National)	Are students in our schools achieving in ways that will allow them to be effective citizens?	Periodic assessment of group mastery of valued targets	Accurate, understandable, usable information about schools and achievement

Source: Adapted from work done by the Assessment Training Institute.

context, standards enable students, teachers, and parents to set goals for educational and personal achievement, and to judge the extent to which those goals have been achieved.

When standards are used as benchmarks, they are made known to students and parents at the beginning of a unit or course of study. Students know exactly what they are working toward. Students are measured against the standard and not compared to each other—thus, it is possible for all students to succeed. Measuring each student against a standard as opposed to comparing them to one another provides for a less competitive environment that encourages cooperation among students.

In most schools today, instruction starts with curriculum standards, which may come from a national organization such as the National Council for the Social Studies, the state department of education, or the local school district. Local education leaders are then responsible for translating these standards into objectives for their schools and classrooms. Once these objectives are in place, teachers and administrators develop and deliver the instruction to help learners achieve the objectives.

Tool for Curriculum Alignment

Pennsylvania has mailed 50,000 resource kits to schools across the state to help teachers align classroom instruction to the state standards. The kit is available on cd-rom and on the state education department's website: www.pde.psu.edu/connections/default.htm.

The explicit nature of the standards is the reason assessment results are such an important part of the standards-based reform efforts that are going on across the country. Through assessments, states, districts, and schools are trying to support and promote the attainment of their standards. To achieve this goal, it is essential that standards, instruction, and assessment instruments are aligned. This consistency is key.

While some would argue that the standards and accountability movement has created a "culture of testing" in many schools today, it would be more appropriate to call it a "culture of assessment" in which there are clear academic standards for what students need to know and be able to do in core subjects. Assessments are the necessary tool to measure student progress toward reaching these standards. In schools that are making effective use of assessment data, the learning improvement cycle starts with the standards, and evolves to assessment, evaluation, and intervention.

Example: Basing Assessment on Learning Objectives

Ms. Bliss is a seventh-grade social studies teacher who is addressing the social studies standard, "People, Places, and Environments," with a unit comparing two states in different geographic areas. One of her objectives for this unit focuses on explaining the difference between the agricultural products of the two states by considering weather factors. Her instruction requires the students to collect and compare average temperatures and rainfall of the two states.

The assessment she has devised measures the behaviors intended by the standard and objective. For example, one of her test items asks the students to identify another state that might produce similar agricultural products based on weather data.

Assessment Data as a Tool for Instructional Improvement

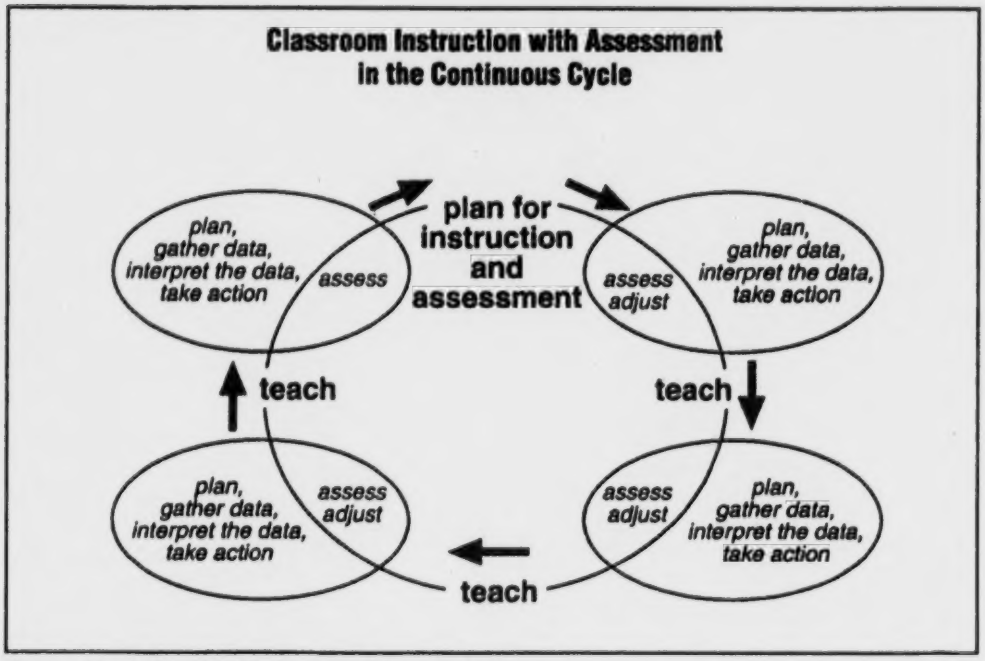
The other trend currently affecting our views of assessment is the realization that assessment results can play a positive, formative role in improving the instructional program overall, as well as the instruction provided to individual students. Brett Gies, Director of Data Analysis and Program Improvement for the Twin Falls School District, has also worked as a teacher, assistant principal, and principal. His understanding of the importance of using assessment data for instructional improvement has evolved over the years. He writes:

> As a classroom teacher, I was frustrated. I didn't always know specifically what my students were learning. And if, as a principal, I asked teachers: "How do you know your students are learning?," they didn't always have a solid answer. That's why I really appreciate the data that are available to us now. If you can't measure learning, you can't manage it. Improvement is left to chance, and a child's education is too important to leave to chance. If we can measure what an individual child knows and doesn't know, then we can provide the instruction he needs. We can manage it. We can improve (Cawelti and Protheroe 2001, 50).

In a general sense, good data are critical to any school improvement effort. As Killion and Bellamy write:

> Without analyzing and discussing data, schools are unlikely to identify and solve the problems that need attention, identify appropriate interventions to solve those problems, or know how they are progressing toward achievement of their goals. Data are the fuel of reform.... In short, using data separates good schools from mediocre schools. Schools that are increasing student achievement, staff productivity and collegiality, and customer satisfaction use data to inform and guide their decisions and actions. Data use essentially sets a course of action and keeps a staff on that course to school improvement and student success (2000).

The North Carolina Department of Public Instruction's *Parents' Guide to Classroom Assessment* (1999) stresses the continuous nature of good evaluation. Teachers should be using assessment data, much of which is generated from tasks embedded in instruction, to constantly "adjust" instruction. This relationship is demonstrated in the following graphic.



Source: North Carolina Department Education 1999.

At the classroom level, regular assessments are important for determining the pacing of instruction, remediation, and introduction of new information. A quiz or other form of assessment can help determine whether the students have grasped an idea and are ready to proceed, or whether the teacher needs to prepare a remedial lesson or additional practice to help them achieve the objective(s).

There are specific characteristics that tell us that assessment is being effectively used to guide instruction in the classroom. These include:

- Assessment is embedded and ongoing and provides prompt, "user friendly" feedback so that adjustments can be made as needed, not just reported at the end of each learning improvement cycle.
- Day-to-day classroom assessments emphasize formative assessments—information that provides an early indication of whether or not learning is taking place—to minimize problems that might arise if learning barriers are not promptly identified and addressed.
- Day-to-day classroom assessments include qualitative data such as information from interviews, journals, and observations as well as quantitative measures within an effective multiple-measures approach.
- Students are told the purpose of an assessment and how this information will be used.

Example: Using Assessment Data to Make Instructional Decisions

After grading the quiz (see page 4), Mr. Jones found that his students were confusing the plot and the setting of the short story. The students needed to understand these two concepts before they started working on their essays. Everyone had already read the same four short stories to start the unit. Two of the stories had a similar plot, and all four took place in different settings. Mr. Jones decided to have the students construct a table where they would summarize the plot and setting of each story as a class activity—and then compare these. Mr. Jones felt the activity would help to clarify the two concepts.

This example from Mr. Jones' classroom illustrates how teachers can use data from assessments to make instructional decisions. Assessments are not an end; rather they are means of improving student performance. Data from assessments—test scores or observations—should be used to determine the next instructional step.

Why Put Time and Effort into Making Effective Use of Assessment Data?

A growing body of evidence from research and practice supports the idea that use of high-quality, targeted assessment data can be a powerful tool to guide instruction.

For example, schools demonstrating success with "closing the gap" and profiled by the North Carolina Department of Public Instruction in a 2000 study were more likely than others to assess students periodically for diagnostic purposes and to disaggregate the data. According to the study report:

> Most of the schools in this study disaggregate data intensely from... districtwide tests as well as from the state's end-of-grade tests, usually with the help of central office staff. Data on all of these tests are reported back to the school in very fine-grained analyses (e.g., by teacher, by student, by curriculum objective, as well as overall results for gender/ethnic groups and grade levels). For example, one school in the study has an early morning remediation session every day where each teacher works with a small group of students on reading and math skills. This school coded their six-week tests so that data could be analyzed separately for each of these small remediation groups. The school used this information to see which teachers were most effective during these sessions (2000, 4).

"Using data to drive improvement" was identified as a key to success in a report developed by the National Education Goals Panel after a series of hearings designed to find examples of successful schools and to understand why those schools were succeeding. Specifically, the successful schools "use performance information to determine where they were succeeding and where they needed to direct their efforts for improvement" (Rothman 2000, i).

In a study of higher-success compared to lower-success Maryland elementary schools, Schafer et al. (undated) found that, in schools they characterized as more successful, "principals are involved with assessment of student improvement and make classroom decisions based on these assessments."

Further evidence was found by Cawelti and Protheroe (2001), who studied six school districts, all of which had significantly increased student performance on state-mandated tests although they served high percentages of students who would typically be considered at-risk. All of the districts that were studied began their improvement efforts by carefully reviewing assessment data.

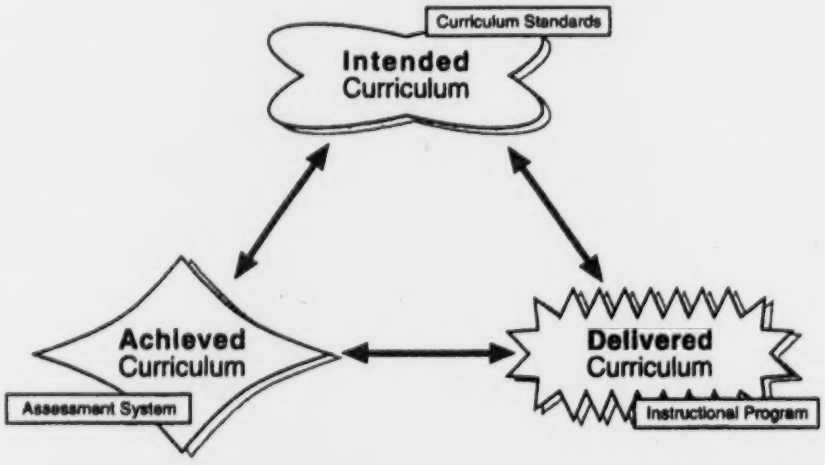
Making It Work: Expectations, Instruction, and Assessment

Although there are many purposes for assessment data and many different decision makers who use this information, an important consideration for schools and districts continues to be the alignment between curriculum, instruction, and assessment. If the assessments are to have an impact on classroom practice—and through that, student learning—understanding and making the necessary alignment of outcomes, activities, and assessment is of the utmost importance.

As Nancy Love writes, assessment is an "integral part of instruction.... The standards talk about curriculum, instruction, and assessment together, as a single vehicle. This reflects a new understanding of the three as an integrated system, not as independent activities. What unites them is a common focus on student learning" (2001, 14-15).

The Alignment Model shown in the box below is used by WestEd and the National Council of Teachers of Mathematics to examine the interrelationships between assessment, standards, and instruction. It illustrates the need for the instruction to provide information directly related to the standards, and for the assessments to accurately measure the intended outcomes. If these three items are aligned, then the teacher, school, and district can focus on school improvement.

Alignment Model



Source: WestEd and National Council of Teachers of Mathematics undated.

Schools continue to be bombarded with societal expectations placed on them by state, district, legal, and other requirements. Available assessments and other instructional materials do not always align with the standards and professional expectations. Thus, the issue of alignment becomes increasingly difficult, but essential, when placed in the context that the major purpose of assessment is to improve student learning and inform instruction.

Notes, Reminders, and Ideas:

Creating Schoolwide Conditions for Effective Use of Assessment Data

Assessment remains distinct from learning in the minds of most American students and teachers.

—American Association for the Advancement of Science undated

If assessment is a critical component of the instructional process, why do so many schools and school districts fail to use it more effectively? While assessment has the potential to improve learning for all students, incorporating it into curriculum and instruction is difficult and complex work. Sound and reflective assessment use takes time—time to learn, and time to practice. The efforts of teachers to use assessment data to guide instruction must be supported with schoolwide conditions and resources. Furthermore, teachers must be given the opportunity to develop the skills they need to assess well and to use assessment information appropriately.

An effective schoolwide or districtwide assessment policy starts with a "Vision of Learning"—a belief that each child can succeed—and a commitment to make that happen. Then the barriers to making assessment much better in serving teaching and learning can be overcome.

Accordingly, principals need to provide for their schools and teachers:

- the **time** to practice their craft;
- the **training** to make it usable;
- the **resources** to support it; and
- a **school culture** that fosters quality assessment practices.

Let's examine four approaches that can help you as a school principal to implement an assessment program that will lead to enhanced student performance.

The Strategies

- Supportive School Climate
- Time for Assessment
- Teacher Training
- Stating Expectations for Performance

Strategy 1: Supportive School Climate

Teacher Beliefs

The school climate must foster teachers' belief in the ability of children to learn and to successfully demonstrate their knowledge and skills on the assessments currently in use. In addition, teachers must believe that the information gleaned from assessments is valid and that it can be used to improve student learning, or a data-driven school culture will not develop (Peterson 1995; Bamburg 1994).

Learning, Assessment, and Grading Communication

Accurate, understandable, and usable information must be provided to students and families about student achievement—principals and teachers must interpret what the scores and grades mean and how they connect to teaching and learning. Families must be educated so that they know what to ask about their child's progress, as well as how to help oversee interventions and improve student outcomes (Northwest Regional Educational Laboratory 1998b).

Priorities

Effectively managing and supporting the ongoing process of designing and developing an integrated program of classroom and standardized assessment must be a high priority. It must be based on an assessment policy that demands and supports quality practices and recognizes the amount of time, resources and money needed to fully implement plans and to assess actions (Aschbacher, Koency, and Schacter 1995).

Continuous and Reflective Use of Data

Interpreting assessment data and then making appropriate changes is an ongoing process. All members of the school community need to be involved in the

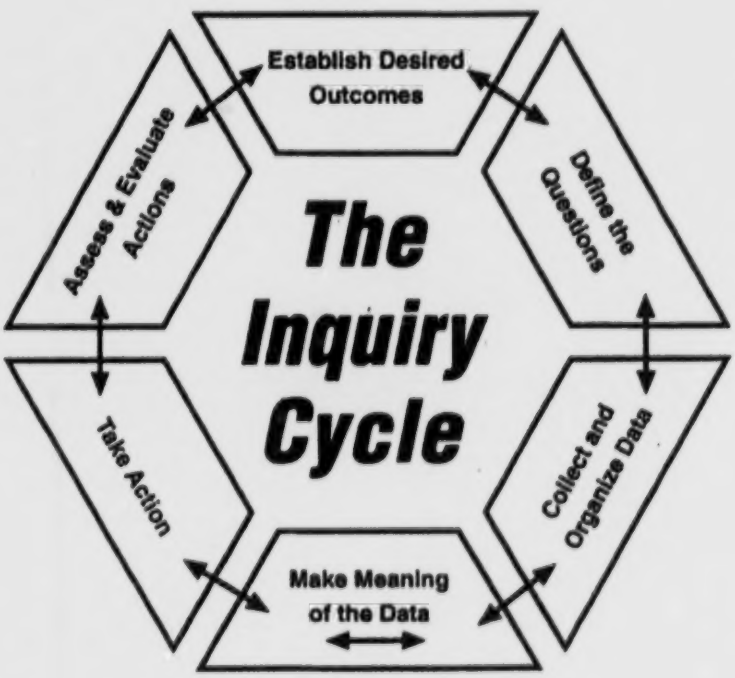
continuous improvement of the school. This cycle involves reviewing the standards or benchmarks, collecting and interpreting appropriate data, determining appropriate actions, and then evaluating the actions (see the figure below).

Leadership

School leadership must develop a vision of excellence in assessment that is in balance with assessment purposes, achievement targets, assessment methods, and ways of communicating results.

The Inquiry Cycle: Continuous, Reflective Use of Data

The Annenberg Institute for School Reform has developed a framework, graphically described below, for this cyclical process of reflection and action that should characterize use of data for school improvement.



Source: Annenburg Institute for School Reform undated. Online: www.annenberginstitute.org/accountability/framework/pgfour.html

Strategy 2: Time for Assessment

Recognition of Assessment as Part of Instruction

Creating appropriate assessments and then using the data they produce takes time, maybe even time away from instruction that teachers are reluctant to give up. Assessment, however, is part of the teaching process; when used effectively, it can make teaching more effective. Through careful planning, teachers can incorporate assessment into their lesson planning process (Aschbacher, Koency, and Schacter 1995).

Synergistic and Practical System

A practical and easy-to-use system is needed to track assessment data and to provide easily accessible reports. If the system is overly complicated or labor-intensive, teachers and students will not use it. Technology can be a useful tool in designing and implementing assessments in the classroom by reducing the teacher's workload. Careful planning of the assessment system can result in an easy-to-use assessment system that provides information teachers need to improve student performance (Rabinowitz and Ananda 2001).

Collection, Storage, Management, and Communication about Student Achievement

Administering a test or assigning an essay is just the beginning of the assessment process. The test, essay, presentation, project, or assignment must be graded. A single test or assignment provides a snapshot of student and class progress in a limited time period. When your assessment data is stored in a larger information system, the results of this single assessment can answer additional questions about student and class progress.

Managing individual student achievement and portfolio information effectively and efficiently requires modern information-processing technologies. An information system can provide you with a summary of an individual student's progress, or picture of the whole class's progress. Similarly, you can compare this year's results to last year's results to determine the benefits of a new curriculum or teaching strategy.

A key component of effective systems is the provision of time on a continuing basis for teachers to discuss the data and to work together to develop "solutions." Teachers view this time as both helping them develop their skills in data analysis and giving them an opportunity to brainstorm and share effective instructional strategies. While common planning time is often difficult to provide, the chart on the following page summarizes Cromey and Hanson's description of the specific systems used in four schools they visited.

Scheduling Approaches for Teacher Collaboration in Four Schools				
	School A	**School B**	**School C**	**School D**
Time and Planning Strategies	1. Once every month, the school day begins two hours later—teachers meet during this time to engage in the activities described below. School makes up this accumulated time by extending the school year.	1. School staff is released early from school once per week for at least 45 minutes. This time is added to other days throughout the week. 2. Entire staff meets once a week for one hour before school. Staff decreases the "nuts and bolts" of the meetings and prioritizes work related to assessment.	1. Same-grade teachers meet informally during weekly planning periods and formally every six weeks. To accommodate these planning periods, students in entire grades are sent to "specials" (e.g., gym, art classes). Time is also allotted at regularly scheduled staff meetings. 2. Teachers are released from teaching duties several days each year and are replaced by substitute teachers. 3. Teachers meet with principal up to three times each year.	1. Teachers request time to meet with each other during school hours; substitutes are hired to support this. In addition, teachers meet after school. 2. Teachers meet in "within-grade" and "subject-area" teams during their planning hours once per week.
Activities	a. School staff rewrites district standards and realigns the assessments it uses accordingly. b. School staff continuously reevaluates this work, and discusses and plans changes as needed.	a. Schools use allotted time to align curriculum across grades and with the state standards. This process is driven by student assessment data. b. School staff continuously reevaluates this work and discusses and plans changes as needed.	a. Staff discusses students' progress according to the "developmental continuums" written by school staff. b. Teachers administer individual assessments to students.	a. Staff shares knowledge gained from professional development activities that addressed curriculum and assessment. b. Staff discusses student mastery of standards and other outcomes and possible intervention strategies.
Source: Cromey and Hanson 2000.				

Professional Collaboration

Teachers need time and opportunities to discuss student work, to develop a common language for discussing challenges and accomplishments, and to develop shared values and transparent criteria for evaluating student work. Working together on the development and use of rubrics will refine each teacher's skills and provide a consistent approach to grading. Similarly, teachers can discuss the development of other assessment instruments to provide both a full range of assessments and to provide for consistent use.

A growing body of work on professional development suggests that talking with one another helps teachers improve their practice and simultaneously work on improving their schools. As with individuals, knowing what works and what does not, figuring out why, and then deciding how to make improvements are essential parts of school progress.

Example: Teacher Collaboration to Develop Assessments

Sue, Bill, Jean, and Renée are a sixth-grade teaching team in a middle school. They teach the core subjects of science, math, social studies, and language arts. Over the past two years, they have worked to implement a broader assessment program that goes beyond paper-and-pencil tests. They are including more real-world projects and realistic assessments in their grading. Students write reports, make presentations, create web pages, and build portfolios.

Jean had recently completed a course on alternative assessments and agreed to teach her team members how to develop and use rubrics. They have developed basic rubrics for written assignments, projects, presentations, and portfolios. Since many of their projects involve two or more subjects, they often grade the materials as a team. This team grading helps them refine their scoring and it gives them an opportunity to discuss individual projects and progress. When they find an area of disagreement, they look for ways to modify their rubric to make the criteria more explicit.

Strategy 3: Teacher Training

Assessment Literacy

Assessment, like other instructional tools, is changing, improving, and evolving. Principals must play a leadership role in ensuring that their faculties are provided and engaged in staff development focused on assessment. Teachers need continual training to enhance and develop their skills. Page 25 contains a literacy self-appraisal that you can use for your teachers and your schoolwide program.

This training is especially important with the shift to using assessment data to improve instruction. The training, however, does not need to be formal workshops or classes; teachers can also examine practices of other teachers and districts to further enhance their own knowledge (U.S. Department of Education 1996).

Reading and Understanding Test Results

Standardized tests are routinely used to assess student and class progress against the educational standards adopted by the district. Principals and teachers need to understand how to read and interpret these test results to make informed decisions. The data from these tests is used to make changes in the curriculum and instructional methods that will help the students master the content related to the standards. Thoughtful interpretation of the results can help teachers focus in on the areas needing improvement rather than doing wholesale revision in an effort to improve achievement.

Coaching and Mentoring

Opportunities must be provided for teachers, individually and collaboratively, to try new practices in a supportive environment with coaching and mentoring.

Strategy 4: Stating Expectations for Performance

Standards

School districts should adopt or create explicit standards and benchmarks *before* developing and implementing an assessment program. These standards and benchmarks communicate expectations to teachers, parents, and students. School administrators and teachers then need to translate these standards into objectives and instructional methods. Assessments are linked to these objectives so that they accurately measure the standards. A lack of connection between the standard, the instruction, and the assessment can severely damage the morale of students, parents, and educators and undermine public confidence in the reforms.

Assessment Literacy Self-Appraisal

	Skilled in ...	Not at all*	Investigating	Emerging	Refining
Teaching-Centric	using assessments to find out whether teaching strategies are working				
	reflecting whether classroom conditions support student learning				
	collecting, using, or adapting successful assessments used by others				
	aligning the written curriculum with actual instruction				
	recognizing inappropriate assessment methods and/or use of assessment information				
	administering, scoring, and interpreting assessment results				
	selecting assessments appropriate for student evaluation				
	developing assessments/assessment tasks with instructional value as well as assessing student learning				
	planning assessment as instruction; not as an afterthought				
	using modern information technologies to analyze and manage student achievement information				
	using assessment results when planning, teaching, and developing curriculum				
	using formative and summative evaluation methods				
Learning-Centric	reviewing/changing assessment instruments and methods if they have cultural, gender, or other bias				
	monitoring student learning regularly, both formally and informally				
	using multiple measures to assess students				
	setting grading scales, establishing standards, determining performance criteria, and developing rubrics				
	evaluating student performance data				
	evaluating assessment results to make decisions about individual student improvement				
	holding students accountable for academic work				
	involving students in peer assessment activities				
	teaching students the scoring systems that will be used to evaluate work				
	communicating assessment results to students, parents, community members, and other educators				
	encouraging parents to monitor student achievement				

*Not at all—I have not engaged in or know nothing about this type of skill; Investigating—I am interested and consciously looking for material on this topic; Emerging—I am beginning to implement this skill and am endeavoring to learn even more; Refining—I regularly use this skill and am looking for ways to perfect my assessment skills.

Some Links to State Standards:

- www.education-world.com/standards/state/index.shtml
- www.statestandards.com/
- whyfiles.org/teach/4.html

Development of Shared Understandings

Teachers need to develop benchmarks for each of the standards. These benchmarks describe student performance that indicates the student has achieved the standard. By developing these benchmarks, teachers have a method to assess student performance on each of the standards. Parents, teachers, and students need to have a shared understanding of what constitutes good work.

Community Involvement

Providing community members with an opportunity to review the standards and assessments can enhance their understanding of the school improvement process and encourage their ownership as a stakeholder. Principals should play a key role in planning and implementing the community engagement.

The chart on page 27 provides an example of an effective way to communicate information about assessment purposes to the community. The chart is a guide designed to help schools communicate with parents and community members about the impact of assessment practices on the learning of young children and the rationale for changes underway in primary classrooms; it serves as thought-provoking material for the seasoned educator as well.

Example: Communicating with the Community about the Purposes of Assessment

Supporting Learning: Assessment/Evaluation/Reporting

We used to...	But...	So now...	Because...
place more emphasis on what students could not or should not do	we learned that this focus undermined the confidence of many students and that we could be more supportive of their accomplishments	we begin with what students can do, then consider their learning needs	this helps them to develop confidence and gives a foundation for building and further refining skills and knowledge
fail students who did not meet pre-set expectations for behavior or ability to do tasks	we found that some students doubted their ability to learn and this increased the probability of their dropping out of school	teachers give students the support needed to allow them to make continuous progress	this maintains their self-esteem and confidence, thus prompting further learning by strengthening the disposition to learn
use pencil/paper tasks as the main way of assessing and evaluating	we now know that this gave a limited view of what students could do	we encourage students to represent their learning in a variety of ways (show what they know)	this provides opportunities for more students to demonstrate their intelligence and to be successful learners
compare learners to each other	this made comparisons more important than the actual learning	each learner is evaluated on what he or she can do in relation to expectations and skills that are widely held as being important, continually refined, and purposefully applied	this helps each student feel valued as a learner and builds on individual strengths, which encourages a good start toward lifelong learning
use checklists for students' report cards	they gave limited information about what students could do	we use information from observations, conferences, and collections of students' work to develop anecdotal reports	they give more comprehensive information about what students can do
[only] use letter grades for reporting students' progress (A, B, C) (G, S, NI)	letter grades were dependent on teacher and parent interpretation and often focused on surface knowledge rather than understanding	we use anecdotal reports to describe students' learning	they give a more detailed picture of what students can do and identify future learning goals
exclude students from the assessment and evaluation process	this did not encourage the development of self-evaluation skills	students are encouraged to take a more active role in assessing and evaluating their own progress and, with the help of the teacher, set future learning goals	as students construct meaning of the world around them, this process encourages self-evaluation, independent learning, and a commitment to further learning
plan conferences for parents and teachers to exchange information	this often overlooked the people with the most relevant information—the students as developing learners	teachers are beginning to plan ways to include students in the conference with parents	together, they can develop a shared understanding of students' abilities, interests, and learning needs, resulting in the setting of realistic learning goals

Source: Ministry of Education, Province of British Columbia 1993.

Notes, Reminders, and Ideas:

The Process of Assessment

The Japanese concept of *kaizen* suggests that quality is achieved through constant, incremental improvement. According to W. Edwards Deming, guru of the Total Quality Management movement, quality in manufacturing is not achieved through end-of-line inspections—by the end of the line, it's too late. Rather, quality is the result of regular inspections (assessments) along the way, followed by needed adjustments based on the information gleaned from the inspections. *Kaizen*, in the context of schools, means ensuring that assessment enhances performance, not simply measures it. (See the Kaizen Institute website for more information about the process: www.kaizen-institute.com/kzn.htm.) In this section, we will discuss the practical steps necessary to produce assessment data that can be used to implement *kaizen*.

Collecting the Data: Summative and Formative Assessment

There are two types of assessment we can use to enhance student performance. First is **formative assessment** that provides immediate feedback to the students to help them learn better. Examples of formative assessments include practice tests and practice essays. Even a pop quiz, given a few days prior to the unit test, can serve as a formative assessment. Various strategies for formative assessments include peer grading and self-reflection. Using the concept of *kaizen*, a teacher would make frequent use of formative assessments to modify and improve instruction.

Second is **summative assessment**. Summative assessments are snapshots of a point in time of a student's learning. Examples include a test given as the unit test and a standardized test given once a year. The purpose of these tests is typically to measure student learning rather than to provide immediate feedback to the student or teacher. Teachers and administrators might use these results to make changes for the next school year.

Formative Assessment

Formative assessments need to be carefully planned to address two goals.

The first goal is to accurately assess student learning. A variety of traditional methods can be used, including multiple-choice tests, written assignments, and projects, as well as nontraditional methods such as games and panel discussions. The traditional methods provide individual feedback and scores, while the nontraditional methods provide immediate feedback, but may not provide individual scores. Another method is to have the students maintain journals that teachers periodically review to provide additional information on individual and class progress.

More Information on Formative Assessment

The following links include information on formative assessment:

- www.qca.org.uk/ca/5-14/afl/sa_formative.asp
- www.fairtest.org/examarts/winter99/k-forma3.html
- arb.nzcer.org.nz/nzcer3/furform.htm

Test questions and activities for formative assessments should parallel those of summative assessment. However, the two should not be identical. For example, a math teacher teaching students how to solve a time/distance problem could create several parallel problems. The formative assessment items should be similar to, but different from, those used on the unit assessment. The names of people and places, modes of transportation, and settings for the problem are easily varied. The task is to teach the students how to solve the problem, not how to substitute one set of time-and-distance data for another.

The second goal of formative assessments is to provide meaningful feedback to the students. One approach that teachers can use to provide feedback is to discuss the items on a test. This discussion goes beyond simply giving the correct answer, but discusses *why* the answer is correct and why others are incorrect. Students should be involved in the discussion to encourage them to reflect on the information. A second approach is peer tutoring, which allows the teacher to pair the students in the class and then have them grade and explain the correct answers. A third approach is to have individuals grade their own work, and then write explanations for the correct and incorrect responses.

The evidence shows that high-quality, formative assessment does have a powerful impact on student learning. Black and Wiliam (1998) report that studies of formative assessment show an effect size on standardized tests of between 0.4 and 0.7, larger than most known educational interventions. Formative assessment is a useful tool to help students make progress towards attaining the standards.

Summative Assessment

Summative assessments are typically given at the end of a unit of instruction and should be used less frequently than formative assessments. Like formative assessments, summative assessments can use paper-and-pencil tests, written work, projects, and presentations. Standardized tests are a form of summative assessment that may be used on an annual basis to measure student achievement of local, state, and/or national standards.

Feedback on a summative assessment varies depending on the type of assessment. A standardized test often provides only a score as feedback. A teacher might also provide only a score on a test or written assignment; but he or she may provide the same level of feedback on a summative assessment as on a formative assessment through remarks or a class discussion of the results.

Looking at the Data and Understanding It

The resulting data from the assessments may be sorted, charted, graphed, or disaggregated—broken down—so that it can be examined from many different perspectives and/or analyzed to discern reasons for a particular outcome. The data can also be compared with other data to look at achievement over time (longitudinal data) as another indicator of student learning.

Teachers, as evaluators, need to be able to identify pronounced patterns in the data so that they can make the necessary classroom or learning accommodations. In this task, an electronic "filing cabinet" that comes with a "built-in statistician" can help make sense out of the data and may be necessary in today's information-laden school setting.

When analyzing assessment data, it is important to ask the right questions. The answers to the questions will help to determine the modifications that are needed to enhance student learning. If you ask the wrong questions, you might change something that is working fine and miss correcting an existing problem.

Reflecting on the Data

Once the assessment data have been organized into a table, chart, or gradebook, educators can interpret the results. What are the strengths and weaknesses of

Questions to Ask

Once data have been collected from the assessments, what questions can educators ask to improve the instruction? Questions that ask "Which, What, Why, or Where?" are better than those that ask "How many?" Asking *which* students are *not* meeting standards in reading is better than asking *how many* students are meeting standards. Examples of useful questions may be:

- How often did the students demonstrate...?
- How did different children (or teams) differ on...?
- Which students/teams di ...? How well did...?
- Where are the gaps? What was disappointing?
- What worked?
- What resources can I/we utilize to address the issues?
- Where do I/we need to change and improve?
- What strategies are most apt to create learning?
- Which strategies need to be abandoned as ineffective?

individual students? What are the strengths and weaknesses of a class? Which strategies worked and which strategies did not work as expected? What could make a difference in how well the students learn?

Reflecting on what individual students know and what they still need to learn can help to determine what to do next (i.e., interventions) to facilitate each student's growth. In addition, the curriculum and/or teaching methods can be evaluated using the assessment results. The assessment system itself can be evaluated to see whether improvements can be made to strengthen the connection between assessment and student success. Individual teachers or the school staff as a whole must use the information gathered to improve relevant aspects of student learning and classroom practices.

The practice in Barbour County, West Virginia, provides an example of detailed assessment data used both to guide instruction for individual students and to shed light on strengths and weaknesses of the instructional program. Elementary school teachers in Barbour County receive an item analysis of the SAT-9 results for the

students they taught the previous year as well as for students currently in their class. Detailed printouts are provided, showing which students did and did not master specific concepts included in the subtests of reading, mathematics, and other subjects.

The chart on page 34 is an example. Using this chart, a teacher would be able to review the individual performance of 22 students, concept by concept.

For example, student #5 demonstrated mastery of all but 5 of the 29 concepts (an x indicates nonmastery). Student #21 had difficulty with 24 out of 29; this student would need more intensive assistance and reteaching. Teachers were also able to use the printouts to identify concepts that most students had mastered—as well as those that were problem areas for many of the students. For example, 80 percent of the students were able to answer items related to rounding decimals correctly, but only 52 percent had demonstrated mastery of rounding percent and converting fractions.

The data not only helps teachers see the specific areas of difficulty for each student, it also helps teachers and principals to pinpoint objectives that either need to be covered more thoroughly or taught in a different way. Teachers can then be given support—staff development, assistance from a master teacher, etc.—with either content or instructional approaches to improve their teaching (Cawelti and Protheroe 2001).

Providing Feedback

Effective teachers provide feedback that focuses on exactly the same issues that the teacher looked at earlier: what the students know and are able to do, what they need to learn to do, and what should be done to intervene. (See page 35 for an example of classroom use of effective feedback.)

Feedback needs to include opportunities to improve and guidance on how to improve. Effective feedback from teachers also provides opportunities for students to express their understanding in classroom dialogues that focus on exploring understanding.

This information is provided to students in ways that enable them to learn better. Bear in mind that the student must understand the feedback to make use of it. To really succeed, however, students must learn to self-assess so that they can understand the main purposes of their learning and thereby grasp what they need to do to achieve. For example, the teacher might have them rewrite the weaker parts of an essay or redo math problems with the wrong answer by using the feedback provided (Black and Wiliam 1998).

Example: Partial SAT-9 Analysis for One Class

ITEM	% of Students Correct	STUDENT # 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Computation/Symbolic Notation																							
Computation/Addition of Mixed Numbers	66	x					x				x				x	x	x					x	
Computation/ Subtraction of Mixed Numbers	33	x			x		x			x	x	x	x		x	x	x		x		x	x	x
Computation/Multiplication of Mixed Numbers	52	x			x		x				x	x			x	x	x				x	x	
Computation/Division of Whole Numbers	57	x			x		x				x				x	x	x				x	x	
Computation/Multiplication of Decimals	61	x			x		x			x	x	x				x						x	
Computation/Division of Decimals	66				x				x		x				x	x	x					x	
Computation/Addition of Integers	80											x			x	x						x	
Computation/Multiplication of Integers	71	x			x						x				x	x						x	
Rounding																							
Rounding Decimals	80				x									x			x					x	
Rounding Fractions	52	x			x				x		x	x			x	x	x					x	x
Rounding Percent and Converting Fractions	52	x	x				x		x	x	x			x	x						x	x	
Computation in Context																							
Computation Context/Subtraction Whole Numbers	61				x	x		x		x					x	x					x	x	
Context/Multiplication of Whole Numbers	76				x			x					x	x	x								
Context/Division of Whole Numbers	52	x			x		x				x	x		x	x		x		x				x
Computation Context/Solving Proportions	42		x		x		x	x			x			x	x	x	x		x	x	x		
Context/Subtraction and Division of Decimals	71				x				x	x					x	x			x				
Context/Addition of Decimals	71	x			x						x				x	x						x	
Context/Subtraction of Decimals	66		x		x								x		x		x				x	x	
Context/Multiplication of Decimals	61	x	x		x		x				x				x		x					x	
Context/Addition of Fractions	61	x			x		x						x		x				x		x	x	
Context/Addition of Mixed Numbers	52	x			x		x				x			x	x	x	x		x			x	
Context/Subtraction of Fractions	38	x			x		x				x		x	x	x	x	x		x		x	x	x
Context/Subtraction of Mixed Numbers	42						x				x	x	x	x		x	x		x	x	x	x	x
Context/Multiplication of Fractions	19	x	x		x	x		x	x	x		x	x	x	x	x			x	x	x	x	x
Context/Multiplication of Mixed Numbers	33	x			x		x	x			x	x	x		x		x		x	x	x	x	x
Context/Division of Fractions	42	x	x		x	x	x				x			x	x		x		x		x	x	
Context/Division of Whole Numbers by Mixed Numbers	61	x	x			x	x							x		x			x		x		
Context/Percent of Decimals	57				x		x		x		x			x	x	x					x	x	
Context/Convert Fraction to Percent	33	x			x	x	x		x		x		x		x	x	x		x		x	x	x

x indicates nonmastery of the item.

Source: Cawelti and Protheroe 2001.

Example: Using Modeling and Frequent Feedback to Improve Performance

Mr. Turner has taught high school chemistry for five years. He is not satisfied with the quality of the student lab reports. He feels that students are not using the lab reports in a mindful way to help them prepare for their exams. At the end of last year, he was frustrated with quality of the reports and the students' performance on the state's achievement test. He had carefully developed his unit to address all the standards on the test, but the students had not developed the higher-level thinking skills they needed.

Mr. Turner decided to try a new tactic this year. The first three units were quite easy and the calculations were simple so that students could concentrate on the reports. He started each lab with an introduction and had the students generate a list of questions they would like to answer in the lab. After the lab, he distributed an example of an excellent lab report from past semesters. He had the students analyze the report and discuss its structure. After the students wrote their lab reports, he graded them and provided lots of feedback. The students then corrected the reports and resubmitted them for a final grading. During the semester, Mr. Turner withdrew the analysis of the reports, but continued to edit and allow students to resubmit the reports.

At the end of the semester, Mr. Turner compared the performance of this class and last year's class. He was very pleased. The scores on the lab reports had increased and the scores on the exams had increased by almost 10 points!

During the next semester, Mr. Turner decided to further wean the students. He agreed to review only two lab reports for each student during a marking period. Although the students complained, he found that the grades remained stable and the first drafts improved with the increased responsibility. At the end of the year, the scores on the achievement test were the best he had seen in his six years of teaching.

The teacher must interpret and communicate standardized test results as well as classroom assessment information. He or she must interpret what the scores and grades mean and then connect the disparate pieces of information to teaching and learning. Using insights gained from this interpretation, the teacher can select new strategies to enhance student learning and progress.

Reporting/Sharing Student Work

Throughout the year, teachers can share examples of high-quality assessments and student work that can serve as a model for students and their parents. The examples also communicate the teacher's expectations to the students and parents. Providing these examples must be done in a lucid way to let students and parents know what variations on such quality look like. Clear scoring guides and examples of student work of varying kinds and degrees of quality are needed.

The information provided to students and parents must be accurate, understandable, and usable. Teachers may also work with parents at this stage to show them how to oversee homework or otherwise assist the student with learning goals at home.

The Tuning Protocol

The Tuning Protocol, developed by the Coalition of Essential Schools (www.essentialschools.org), is a structured, facilitated conversation that involves a group of teachers and sometimes others in examining samples of student work to address a focusing question—usually about instructional practice or the quality of student work. It is often used as a way to address questions of goals or standards for an individual classroom or school.

A typical tuning protocol brings together a small group of teachers (10-15) with a facilitator, who may come from outside or inside the school. One teacher or a team of teachers presents sample student work and the context for the work (the assignment or assessment rubric). The protocol follows a schedule calling for presentation of the task or project; clarifying questions; examination of student work samples; "warm" and "cool" feedback to the presenting teacher; reflection by the presenting teacher; and debrief of the process. The entire protocol takes about an hour and 15 minutes.

The protocol may be used as a "one-shot" instrument for addressing an important schoolwide question or issue or, probably more effectively, as part of ongoing professional development or planning. The Tuning Protocol can be modified and adjusted for use in the classroom as a peer review and exhibition tool. For more information, go to www.essentialschools.org/fieldbook/schdesign/inquiry/examstudwork/caprotocol.html.

Guidelines for Looking at Student Work

The following guidelines were developed by the Philadelphia Education Fund to help teacher teams get started with a process of inquiry. These guidelines were informed by the work of a variety of education organizations and practitioners.

1. Gather a team, or small group, of teachers together.
2. Select a piece of student work. The sample should demonstrate a rich variety of student learning. It can be a work-in-progress, a final piece, or a document of a performance. Also collect the scoring guide or rubric used to assess that piece. Make copies for team members, if possible.
3. If someone in the group is not familiar with your unit of study, take a few minutes to introduce its overall purpose, the activities that have been conducted, and the work that has been generated.
4. Discuss and write down one standard from those that you expected students to address in this activity. What did you expect the students to know and be able to do?
5. Next, take a few minutes to look at the work as a group. Either read it aloud, or let each person take a turn looking at it.
6. Write down the group's observations about the work. Then write down comments and questions. You might allow each team member to do this first individually and then share in turn.
7. Next, use your scoring guide or rubric to assess the piece of work. If you do not yet have a scoring guide, reread the standard you have identified and assess the work based on its criteria. (You might put together an informal rubric by doing this.)
8. Take a few minutes to discuss as a team the following questions: What can you see from your observations, comments, and questions that will help you assess student learning? How might these observations determine your next steps as a teacher? Do these observations tell you anything new about your unit of study or classroom activities?

Source: Philadelphia Education Fund undated.

Intervention: Taking Action Based on the Data

While use of assessment data can help to identify students who are not mastering particular objectives—or who are generally below grade level—that knowledge is meaningless without providing supports for these students. In their study of "high-performance districts," Cawelti and Protheroe found that a common characteristic was the districts' recognition of the need for:

> instructional processes that enable teachers to accomplish three things on a daily and weekly basis: (1) organizing instruction to regularly administer interim assessments of skills taught before moving on to new material, (2) providing tutoring or extra help for those students who fail to master the *skills taught and enrichment learning activities for those who have mastered the skills,* [emphasis added] and (3) providing frequent practice throughout the year to ensure retention for students who have initially mastered the skills needed (2001, 98).

The provision of these services was handled in a variety of ways—through tutorials, after-school programs, summer school, etc. Often, it required very detailed efforts to create school and grade level schedules that made it possible to group students by skill level for parts of every day. Instruction is typically provided in small groups, with opportunities available to help teachers refine the skills they need to work with students who need additional help to achieve mastery. The efforts required leadership on the part of the school principal, commitment from the entire school staff, and a highly organized and flexible system—but results in the form of higher student achievement proved the value of the efforts. Assessment empowers teachers and students to modify, customize, and improve practices as needed. As data is gathered, it is possible to determine changes and to adapt and adjust teaching and learning accordingly—congruent with classroom instructional practices.

The best approach is to begin with a list of needs generated during the evaluation process and then to brainstorm possible responses for each of the areas needing attention. Then take action to improve student learning based on what the data reveal:

- Decide what modifications are needed in day-to-day instruction or practice and how to change them.
- Identify strategies and resources to help students improve performance.
- Translate those strategies into good instructional and assessment practices and use them daily.

Picture effective coaches and sponsors of drama clubs as they involve their students in scrimmages or dress rehearsals. Such activities serve to identify problems and weaknesses, followed by more coaching and opportunities to practice or revise.

The ongoing interplay between assessment and instruction, so common in the arts and athletics, is also evident in classrooms using practices such as non-graded quizzes and practice tests, the writing process, formative performance tasks, review of drafts, and peer response groups. The teachers in such classrooms recognize that ongoing assessments provide feedback that enhances their instruction and guides student revision (McTighe 1996).

What Does an Intervention Look Like?

Interventions are methods used to engage the learner with the content that was not mastered in the regular instruction, with the purpose of enhancing understanding and learning. Interventions should not just repeat the same instructional methods that were not effective with this specific learner before. Teachers may explicitly teach a particular strategy to a group of children, offer manipulatives, introduce mnemonics, provide graphic organizers, webbing and mapping techniques or other scaffolding devices, change the delivery method, change the grouping arrangement, use non-graded quizzes and practice tests, set up peer response groups or teach the writing process. The following are just a few examples.

Writing Interventions. Students usually need help learning how to write analyses of academic contents. A teacher can provide pieces of exemplary writing from the previous year, have students share their own writing, give students written suggestions for improving particular pieces, or involve students in the creation of a rubric for evaluating writing. The scenario from Mr. Turner's chemistry class on page 35 is an example of a writing intervention he used to improve both the writing of the reports and test performance. All of these experiences help students come to understand what high quality content-area writing is. Together, these strategies form a supportive web of interventions for those students struggling with the process of learning to do written assessments.

Extension Interventions. Helping students to identify and build on strengths is another part of the intervention process. For example, a student who writes well can use that skill to help organize an oral presentation. A student who draws well can use a series of sketches to structure a storyboard that will help to scaffold his less-skilled writing.

Assessment Interventions. Assessment can be turned into an instructional intervention in many cases by simply involving students in the assessment and evaluative process. Four of the most common ways of involving students are through self-assessment, peer assessment, learning contracts, and student-led conferences or parent/student/teacher conferences.

The key to a successful assessment intervention is actively engaging the learner in processing the feedback. For example, Mr. Turner involved his students with the

process by having them rewrite their lab reports using his feedback. In prior semesters, he found that students did not use his feedback in a mindful way, because they had already started a new unit of instruction when they received it.

Outside Interventions. Not all interventions are instructional. A needed intervention might be a referral—for example, requesting a hearing test, referring a student for diagnostic testing, or asking a site specialist for additional resources or help.

Using Assessment Data to Identify Good Teaching

A 15-year study at the University of Tennessee-Knoxville found that the quality of teaching in a student's past accounted for differences in standardized test scores of as much as 50 or 60 percentile points. Furthermore, effective teachers can't make up for ineffective teachers—the effects of a bad teacher do not go away over time. The study shows that children who have had two consistently ineffective teachers do not perform up to the achievement level they would have if they had been assigned to better teachers.

The recognition of the difference that teacher quality makes to a child's cumulative learning underscores the importance of having an accurate way to measure teacher quality. Assessment data can make a useful contribution to this effort. However, it is crucial that such data be used effectively and appropriately.

Resources for Accountability

The Tools for Accountability project collects, evaluates, and disseminates information about a wide variety of "tools" that help account for and improve teaching and learning. In its "toolbox" are such items as:

- Sensible ways of analyzing standardized test scores to improve instruction and monitor quality.
- Instructive guides that help teachers and parents to critically examine samples of student work.
- Practical activities that help parents better understand their children's progress in learning.

Source: Annenberg Institute for School Reform undated. Online: www.annenberginstitute.org/accountability/toolbox/default.html

While annual norm-referenced, standardized tests are presumably used to determine how well the school has taught the child during the preceding year, the tests actually determine the cumulative level of knowledge of students—not what knowledge was added during the year.

Since the real measure of whether instruction is successful is how much *progress* a student makes in learning, a data-driven school may want to investigate the *Value-Added Assessment System*—a method pioneered by William Sanders in Tennessee for measuring student academic progress relative to each student's starting point. (For a brief description, see Sanders' 1998 article "Value-Added Assessment," posted at www.aasa.org/publications/sa/1998_12/sanders.htm.)

Notes, Reminders, and Ideas:

Enhancing Classroom Assessment

When you are shopping for a car, you probably consider a number of factors. The style and color of the car may be high on your list; you probably look at the gas mileage, the location of the dealer for service, and the price. Finally, you take the car for a thorough test drive.

Similarly, good educational assessments require more than a single paper-and-pencil test. In this chapter we will consider two concepts that can enhance assessment efforts and make them more meaningful.

- First, we will examine the use of **authentic assessment.**
- Second, we will consider how **multiple assessments** can provide a better picture of student progress than a single measurement.

Authentic Assessment

We have all experienced tests that essentially measured what we could memorize the day before the test and will forget the day after the test. Tests that required us to correctly spell 15 words that were unknown to us three days ago, to name the states that comprise the southwestern part of the country, and to produce the formula for calculating the area of a triangle all measure what we memorized *for the test.* Spelling the 15 words correctly or writing down the formula for calculating the area of a triangle has little to do with actually *using* that information in the real world. In a sense, it is sterile information, void of any real-world context.

Authentic assessment, sometimes referred to as *alternative* assessment, embeds the knowledge or skills to be assessed in a real-world context. Often, the assessment requires the student to engage in problem solving to generate a correct answer. For example, rather than having the teacher read the words and the student write them down for a spelling test, the assessment might require the student to design a poster

or story incorporating the words so that they are correctly used and spelled. Similarly, the student might be given the task of calculating how much paint is needed for a stairway where part of the wall forms a triangle.

What are the advantages of authentic assessments? Such assessments create a stimulating environment that requires students to go beyond memorization. Students must apply their knowledge in a realistic setting. This application develops useful knowledge and helps students understand how they can use what they have learned in the real world.

Creating Authentic Assessments

Authentic assessments require consideration of how to use the information or skill in a realistic setting. A good starting point is to identify how the student can use the objectives in the real world. Once an application is identified, the task is one of embedding the knowledge or skill application in a realistic context. Thus, rather than asking the worn out question of when the cars leaving San Diego and New York will meet, the teacher may create a context for the time-distance problem using local landmarks (for example: "Your friend calls you from the Cherrydale library and wants to meets you at Ayers Hardware Store."). The same process is used to solve a problem dealing with two people walking a block or for two spacecraft traversing the universe.

More Information on Authentic Assessment

The following URLs provide brief background information on authentic assessment:

- www.teachervision.com/lesson-plans/lesson-4911.html
- www.eduplace.com/rdg/res/litass/auth.html

Multiple Measures of Performance

> ***Teachers, parents and students develop an understanding of what is valued in education by virtue of what is assessed and how it is assessed. For example, when multiple-choice tests are used exclusively, it gives the impression that there is only one right answer to every question and that there is always a right answer.***
>
> —Northwest Regional Educational Laboratory, 1998.

Paper-and-pencil tests are only one part of a balanced and effective approach to assessment. As the purposes of testing differ, so do the types of testing required to

measure student achievement. A multiple-measures approach using different types of tests is often needed to accurately and fairly assess student progress.

This approach puts the appropriate type of assessment to work for the right purpose. Performance assessments, for example, might be used for instructional purposes, while norm-referenced tests are used to generate comparative information. Similarly, a spelling test is used to determine whether students can correctly spell a set of words, while a portfolio is used to show progress in writing skills over an academic year. The key is using the assessment that best provides the desired information for an assessment-literate school culture.

Example: Using Multiple Measures of Performance

Mrs. Carr's sixth-grade students are studying a unit on forestry that is inquiry-based and lasts approximately two weeks. She will use several different forms of assessment to determine student progress.

At the beginning of the unit, the students start a KWL chart by listing what they know about forest products and what they want to know. These questions guide their inquiry and research.

After four days of work, Mrs. Carr wants to check the students' understanding of some key concepts. She creates a list of 10 true and false questions that she asks the class. Students indicate true with thumbs up and false with thumbs down. Two days later, the students complete a 15-item test. At the end of the unit, the students will submit a report on their research and do a PowerPoint presentation. Both of these assignments will be graded using rubrics. On the last day of the class, the students complete the KWL chart by listing what they have learned.

Mrs. Carr is using a variety of assessment instruments in her class. She only uses one traditional method—the 15-item test. The others—the KWL chart, the thumbs up or down test, and the rubrics for the report and presentation—are all examples of alternative assessments.

Carefully considering all the types of behaviors to assess and the reasons to assess them, and then selecting the assessment approach that best matches these targets and purposes, creates a balanced approach for the classroom. Sometimes the

answer will be an alternative assessment, sometimes a traditional assessment, and sometimes a combination of the two. It is important to remember that:

- Differing forms of testing can, if used correctly, enrich the understanding of student achievement.
- The assessments we choose must work synergistically.

If we wish to take full advantage of the power of assessment to maximize student achievement and meet the information needs of those who seek to promote excellence from the classroom to the boardroom, we must know what we want to accomplish with an assessment. That question then leads to what assessment techniques to use and when to use them. The answer may be a combination of high-quality standardized tests and high-quality classroom assessments used at varying intervals..

Stiggins (1997) suggests that a teacher first determine which achievement target to assess and then select an appropriate assessment method(s) to assess the target. The chart on page 46 shows this process. This chart illustrates that no single assessment is appropriate for all assessment activities. Any one type of test is only one part of a balanced approach to assessment.

Accountability-minded schools and teachers recognize that they can only be sure about an outcome or perception if it is corroborated by multiple forms of evidence. They seek multiple measures to provide a rich portrayal of student learning. This "multiple-measures approach" to assessment is the key to obtaining valid, reliable, fair information about student achievement.

What About the Standardized Test My District Gives Every Year?

> ***The model of the external, multiple-choice, norm-referenced test dominates the underlying conceptions of what it means to assess and how to do i t.***
>
> —Neill, undated

Standardized assessments developed by test publishers—the type of test that is commonly used to evaluate student learning over time in comparison with others—usually receive the most attention. Such tests were not designed to provide immediate feedback on student progress needed by students, teachers, or parents. They are typically used only once a year, and provide objective information about each student's progress in mastering knowledge and skills.

Educators and the public have traditionally perceived standardized tests as exclusively norm-referenced, multiple-choice examinations, but that is not true today. A standardized test is one that is always given in a consistent manner, with the same

Aligning Assessment Methods with Purposes

Target to be Assessed	METHOD			
	Selected Response	**Essay**	**Performance Assessment**	**Personal Communication**
Knowledge mastery	Multiple choice, true/ false, matching, fill-in can sample mastery of elements of knowledge.	Essay exercises can tap understanding of relationships among elements of knowledge	Not a good choice for this target—three other options preferred	Can ask questions, evaluate answers, and infer mastery, but a time-consuming option
Reasoning Proficiency	Can assess application of some patterns of reasoning	Written descriptions of complex problem solutions can provide a window into reasoning proficiency	Can watch students solve some problems or examine some products and infer about reasoning proficiency	Can ask students to "think aloud" or can ask follow-up questions to probe reasoning
Skills	Can assess mastery of the knowledge prerequisites to skillful performance, but cannot rely on these to tap the skill itself	Can assess mastery of the knowledge prerequisites to skillful performance, but cannot rely on these to tap the skill itself	Can observe and evaluate skills as they are being performed	Strong match when skill is oral communication proficiency; also can assess mastery of knowledge prerequisite to skillful performance
Ability to Create Products	Can assess mastery of the knowledge prerequisite to the ability to create quality products, but cannot use these to assess the quality of products themselves	Can assess mastery of the knowledge prerequisite to the ability to create quality products, but cannot use these to assess the quality of products themselves	Can assess: 1) proficiency in carrying out steps in product development, and 2) attributes of the product itself	Can probe procedural knowledge and knowledge of attributes of quality products, but not product quality
Dispositions	Selected response questionnaire items can tap student feelings	Open-ended questionnaire items can probe dispositions	Can infer dispositions from behavior and products	Can talk with students about their feelings

Source: Based on Stiggins 1997.

directions, the same questions, and the same time limits. Thus, scores can be compared with confidence in test validity and reliability. All assessments administered within a state or local testing program should be standardized, no matter what type: performance based, norm-referenced, or criterion (standards) referenced.

Standardized Tests vs. Diagnostic Assessments

Assessment specialists explain the difference between standardized tests and diagnostic assessments by pointing to mountains:

A standardized test compares mountain climbers to each other, but a diagnostic assessment tells the climber and their coaches where they are on the mountain (Jones 1998).

Assessment Toolbox

While assessment is one of the most important pieces of work teachers do, it is the one role they are least prepared for.

—Beth Bader, AFT

Teachers and administrators must have an "assessment toolbox" that is full of balanced and effective high-quality instruments—both standardized tests and classroom assessments—and these instruments must be in the hands of an assessment-literate culture. How well stocked is your toolbox? See page 25 for a self-rating scale that could be used by individual staff members or as part of a discussion focused on what the school needs to do to improve its overall assessment literacy.

Using a variety of assessment techniques can provide a rich set of data that can provide insights into your teaching and students' progress. With some planning and creative thinking, you can develop a variety of assessment tools that will help you modify and improve the learning environment in your classroom.

Notes, Reminders, and Ideas:

The Continuous Improvement Process

Schools have traditionally been slow to change and to adopt new ideas. In the past, we have often seen schools implement a reading or math program and then tend to ignore it until the end of the school year. Achievement scores or final grades were evaluated and a decision was made to either continue with the program, implement a new program, or revert to an older method.

The increased emphasis on and use of standards and benchmarks, however, has created a new environment. This environment provides a large amount of data throughout the school year—data that is current on a daily, weekly, or monthly basis. Using this data, teachers and administrators can now make minor adjustments in the teaching strategies and school environment to enhance student achievement in relation to the objectives. In this chapter, we will focus on a process to continually improve student achievement.

Achieving ever-higher levels of classroom or school performance requires a well-executed approach to continuous improvement. The principal plays the key role in creating the climate for continuous improvement. The continuous improvement process has several important characteristics and needs to be "embedded" in the way the school operates:

- Communicate clear achievement and performance targets for the students, staff, community, and school.
- Use collaboration and data-based inquiry to improve performance.
- Include cycles of planning, execution, and evaluation.

Continuous improvement must become a dominant characteristic in the school culture. This process is based on the assumption that schools can continually improve

methods, materials, people, and environment to achieve better results—to improve performance by examining the practices of the school and making incremental improvement plans. Questioning, seeking data to inform those questions, reflecting, and implementing informed action plans are simply "the way things are done." Continuous improvement is applied throughout the school year to make adjustments to enhance learning. Teacher-designed assessments, rather than just an end-of-the-year standardized test, provide the data for the continuous improvement process.

For all students to achieve the highest standards, the school's staff must foster an environment that encourages and uses data-based inquiry and builds a strong foundation for continuous learning and discovery. Total Quality Management theory and practice is one avenue to enable educators to start down the path of continuous improvement. The collection and analysis of data, a cornerstone in the W. Edwards Deming Total Quality Movement that is widely used in industry, is rare in many school districts.

There are three major steps in the continuous improvement process:

- First is the specification and communication of target performances.
- Second is the collaborative use of data to determine whether the standards are met.
- Third is the continuous improvement cycle. The following sections will describe each of these steps.

Communicating Achievement and Performance Targets

In today's information age, all students need to meet high standards to succeed. The standard is no longer "doing better than the rest," but rather meeting a previously agreed-upon, rigorous level of knowledge and skills. The people, the vision, and the work of the school must all focus on these clearly communicated and defined standards.

All participants must discuss and define criteria that can be used to identify satisfactory outcomes. They must set standards for judging the quality or degree of success toward these outcomes. They must also affirm their obligation to be accountable for particular outcomes. They must ask:

- What do we want children to know and be able to do?
- What exactly are we saying we are responsible for, and what standards will we use to judge our success?

Standards define curriculum and instruction. Administrators, teachers, and parents must invest as much time and energy on achieving standards as on setting

them. Without an accountability architecture, standards are only "words on a page, not engines for reform" (Pimentel and Arsht 1998).

To meet these high standards, teachers and administrators must create a place where everyone experiences powerful learning and teaching. Classroom and organizational practices and structures must be in place that support and encourage all students to reach high standards. Each student must have the opportunity to learn to the same high standards across the school. Curriculum and instructional strategies must promote students' mastery of the skills, knowledge, and attitudes.

Multiple forms of assessment, aligned to the curriculum, must be used to measure how well all students meet the standards. Provisions must be in place to support student achievement. Benchmarks must be set for judging the quality or degree of success toward these defined outcomes. Assessment and accountability data must be regularly collected, disaggregated, and analyzed for the purpose of continuous improvement. School must be a place where students, teachers, principals, and staff build a foundation for continuous, innovative learning and the school continuously strives to improve (Cross 1998).

Assessments define standards. Assessments are the final word in defining what educators and the community want students to know and do. Regardless of what anyone says in their standards documents (or in course outlines), the assessments define what was really meant. So, assessments communicate the standards to which school systems, schools, teachers, and students aspire. Teachers have an obligation to ensure that their assessments correctly and adequately measure the objectives related to the standards.

Using Data to Determine Whether Standards Have Been Met

In the data-based school, decisions to make a change or modification in instructional strategies, assessment, or the school environment are driven by assessment data. Teachers and administrators review teacher assessments as well as standardized testing data to determine how best to improve student learning. Using this data-based approach can focus efforts to enhance learning on the weakest parts of the instructional environment.

The Framework for Accountability developed by the Annenberg Institute for School Reform (www.annenberginstitute.org/accountability) and the Instructional Improvement Through Inquiry and Collaboration (IITIC) project of the Coalition for Essential Schools (www.essentialschools.org/fieldbook/iitic/iitic.html) have laid the foundation to help schools use data-based inquiry to address schoolwide challenges. Each school will move through the following phases and questions in a way that best fits their own goals, data, context, and challenges.

Phase 1: Assemble and analyze data. During this part of the process, participants articulate the questions they have about the school. For example, a group of teachers might ask:

- What do we want to know about our teaching? About our students?
- Is something happening that pleases or bothers parents?
- How will we know whether we are doing something well or poorly?
- Have we considered hypotheses based on student "output data" (such as test scores, grades, etc.) as well as "school input data" (such as efficacy of instruction, strength of curriculum, cultural awareness of students' context and backgrounds, etc.)?

Other members of the school community may have different questions. The questions chool-board members might ask about overall performance often differ from questions asked by middle school teachers. These, in turn, may differ from those posed by parents about graduation requirements. The state or district may question standardized scores. Taxpayers ask whether programs are working.

- What do parents say is our challenge? Students? Community members?
- How is this challenge or question related to our schoolwide goals for high achievement and equity?
- What are our hypotheses about why this challenge exists?
- How could we collect or draw upon data to test these hypotheses? How could we utilize research literature in testing these hypotheses?
- Whom might we need to ask in order to get ideas for hypotheses?
- Can we identify one hypothesis that is most supported by data?

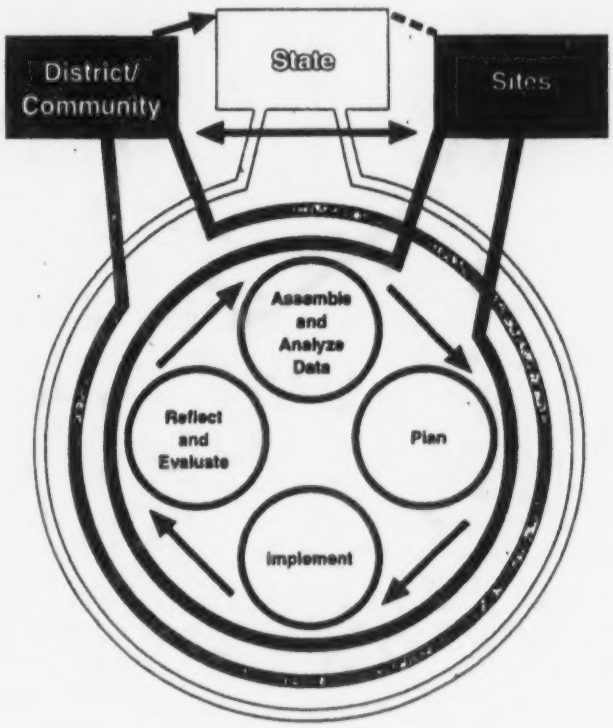
While all questions are brought to the table, the school community must set priorities for what it can and must address at any given point. A significant part of this activity is determining a specific focus for the next steps in the inquiry.

Focus questions frequently address outcomes (e.g., attendance or dropout rates). Or, the questions may call for descriptions of behaviors that might lead to those outcomes (e.g., the types of assignments teachers give to students; the GPAs and standardized test scores of students who are absent more than 10 percent of the time; the correlation between the mathematics program and scores on the SAT).

When the questions address outcomes, participants need to take the further step of deciding criteria to judge the level of success. For example, they might ask: What dropout rate would satisfy us that we are doing an adequate job?

Guiding Assumptions of Data-Based Decision-Making

- Data based decision-making focuses on improving student results for all learners and is an integral part of the continuous improvement process.
- Decision-making supports opportunities for all learners in meeting standards.
- The continuous improvement process is guided by quantitative and qualitative data drawn from the classroom, school site, district, and statewide sources.
- Stakeholders draw on an understanding of learning to select improvement strategies.
- What do we want to accomplish?
- By whom? By when?
- How will we know we are done? Quantity? Quality?
- Change is initiated.
- Improvement strategies are evaluated.
- Efforts are refocused as to what more needs to be done to meet goals based on what the measurements say.



Source: Minnesota Department of Children, Families, and Learning undated.

Phase 2: Collect and organize data. Once the focus questions are defined, we can begin the process of accumulating the evidence needed to answer them. Much of the data may already exist in various forms. Some forms are externally mandated; some are already recorded routinely; sometimes participants just need to locate them and organize them for use. Participants may also choose to generate new data (e.g., surveys, work samples). To answer their focus questions, participants often need to consider multiple measures and to disaggregate the data by specific groups.

Survey instruments can offer reliable measures of a school's efforts to meet students' academic and developmental needs. They can collect data on attitudes, practices, and overall climate and can be tracked longitudinally (year to year). Survey results often augment other data collection/analysis methods.

Some important questions to ask are:

- What kinds of evidence will suit our questions (e.g., test scores, reports, portfolios, survey data, anecdotes, or behaviors)?
- Where can we find the data?
- Will we want to be able to disaggregate the data by groups?
- How can we collect the data?
- How can we store information so that it is easily accessible?

Accountability-minded schools recognize that they can only be sure about an outcome or perception if it is corroborated by multiple forms of evidence. This means, for example, supplementing data from multiple-choice standardized tests with performance assessments, the examination of student work samples, or observations.

Basic Data Elements

Basic Data Elements for Elementary and Secondary Education Information Systems can be downloaded from the U.S. Department of Education's website. This document contains a set of basic student and staff data elements recommended by the Core Data Task Force of the National Forum on Education Statistics. The guide is intended to provide a common language to promote the collection and reporting of comparable education data. The report also contains a process for identifying and updating the set of data elements to be maintained by a school, school district, state education agency, or other education unit with a need for student and staff information.
See nces.ed.gov/pubs97/97531.pdf.

To help the school, parents might volunteer to assist in data collection, or businesses might lend support or resources. The collection of data goes far beyond standardized tests and often includes the entire range of school stakeholders.

Finally, the school makes decisions—for example, acquiring an especially designed software package—that facilitate the organization and management of data.

Make meaning of the data. Individual pieces of data are of little if any value in their raw form. They are simply bits and pieces—words and numbers—waiting to be organized into meaningful patterns. Participants in this activity conduct close readings of all materials, including both numbers and narrative data. They group data according to the focus questions in order to articulate relationships, categorize, or note patterns and rules. They interpret written as well as quantitative analyses. Participants look for unexpected outcomes and relevant models or theories that might clarify ambiguities. In sum, they seek to assign meaning to the data.

During this process, data are transformed into information that may be used to improve practice: how our students are doing in the reading curriculum; which students take advanced courses; which students are absent most often. Information—patterned data—is what accountability-minded schools use to make choices, guide decisions, and shape policy. The questions we ask about the data are the key to identifying patterns. It is the school's purposes and questions that turn data into meaningful information.

The Continuous Improvement Cycle

The next three steps—plan, implement, and reflect and evaluate—represent the improvement cycle. These steps are described in the following pages and are best facilitated by the principal and/or other instructional leaders. An excellent example of a school district's improvement process based on these elements is described on pages 56-58.

Plan. This step involves school staff collaborating to brainstorm potential solutions. They might address the following questions:

- Based on our focus questions and the hypotheses we have identified, what are some solutions that might help?
- What ideas or expertise exist in the school that could help provide solutions?
- What research literature could provide other solutions?
- Are there other schools that are implementing exciting ideas relevant to our challenge? How could we learn from them?
- Whom might we need to ask in order to get ideas for solutions?

The Brazosport 8-Step Instructional Process

The Brazosport Independent School District in Texas has developed and uses the 8-Step Instructional Process, a Plan-Do-Check-Act approach of TQM customized for a school setting. Past-superintendent Gerald Anderson characterizes the process as critical because it:

- prioritizes what's important;
- focuses on these issues;
- guides data analysis; and
- provides maintenance and tutorials that sustain learning over a long period.

This key to Brazosport's success was first identified when the district used 1993 TAAS data to answer the questions: Which Brazosport teachers are already accomplishing high levels of student achievement with all students, especially economically disadvantaged and minority students, in their classes? And how are they doing it?

One of the teachers identified was a third-grade teacher who was only in her third year of teaching. As she explained the approaches she used, it became obvious that she had embedded the use of assessment data in her instructional planning. Specifically, assessment data helped her focus on meeting the needs of every student in her classroom.

The district asked the teacher to collaborate with other teachers to develop a more structured model that could be used for staff development. She then trained teachers and other staff in Velasco Elementary School using the model. In essence, the process begins by using assessment data to identify what students, either as a group or individually, already know and what they still need to work on. Teachers then develop a campus timeline that encompasses all objective areas and time allocations based on student needs. Students are periodically reassessed, with either tutorials or enrichment provided as necessary. Periodic reteaching of objectives ensures that students retain what they have learned.

Significant improvements in student test scores at Velasco in under a year confirmed the potential of the model, and the teacher was asked to leave her classroom and provide districtwide training and support for implementation of the model.

The model, as it currently exists in Brazosport, has evolved into more than a structure. It is a state of mind embedded in the daily life of teachers and schools. Data drives what instruction is provided, to whom, when, and how.

Included below are details about the process (also see page 58):

1. **Disaggregate data:** Teachers receive individual student and classroom reports on TAAS results for both their previous year's classroom and the students they will teach during the current year. Principals receive the same information. Data are also available from periodic assessments developed by teachers in the district.
2. **Develop a timeline:** Based on both the knowledge and skills in the Texas Essentials Knowledge and Skills and the assessment data available from the TAAS, grade-level teachers develop an objective-based teaching calendar for the year.
3. **Deliver instructional focus:** The timeline is used by each teacher to identify the objective to be taught. The instructional focus—the objective—is announced and taught at the beginning of each day or class period.

The Brazosport 8-Step Instructional Process

4. **Administer an assessment:** Assessments, some commercially developed and many developed by district teachers, are administered periodically to ensure that students have mastered the objectives taught during the specified time period. Data from these are used to determine whether whole classes need additional reteaching or if special assistance is needed for specific students. The approach is intended to help teachers detect and correct problems early.

5. **Based on assessment, deliver to students either enrichment activities or tutorials to reteach:** Enrichment activities are provided for those students whose short-term assessment shows they have mastered the skills just taught. They are scheduled during the time other students are involved with tutorials.

6. **Use tutorials to reteach.** Tutorials are provided to students who have not mastered the objectives just assessed. The tutorials typically include fewer students than a regular class. Help is also provided after school and on Saturdays.

7. **Maintain and reteach:** Teachers include short, periodic reteaching/maintenance activities in their instruction to ensure that skills already taught are retained.

8. **Monitor:** Principals visit classrooms during the time allotted for the instructional focus to monitor progress and to maintain knowledge of the progress of individual students and classes.

Charlene Burgner, Director of Instruction, emphasizes the importance of how the schools use test data to improving instruction:

> The TAAS has become *part* of our teaching strategy, and I mean that in a good way. We don't teach classes—we teach to a child. And we can do that since the principals and teachers know the specifics of how each child is doing based on TAAS results or on other assessments.

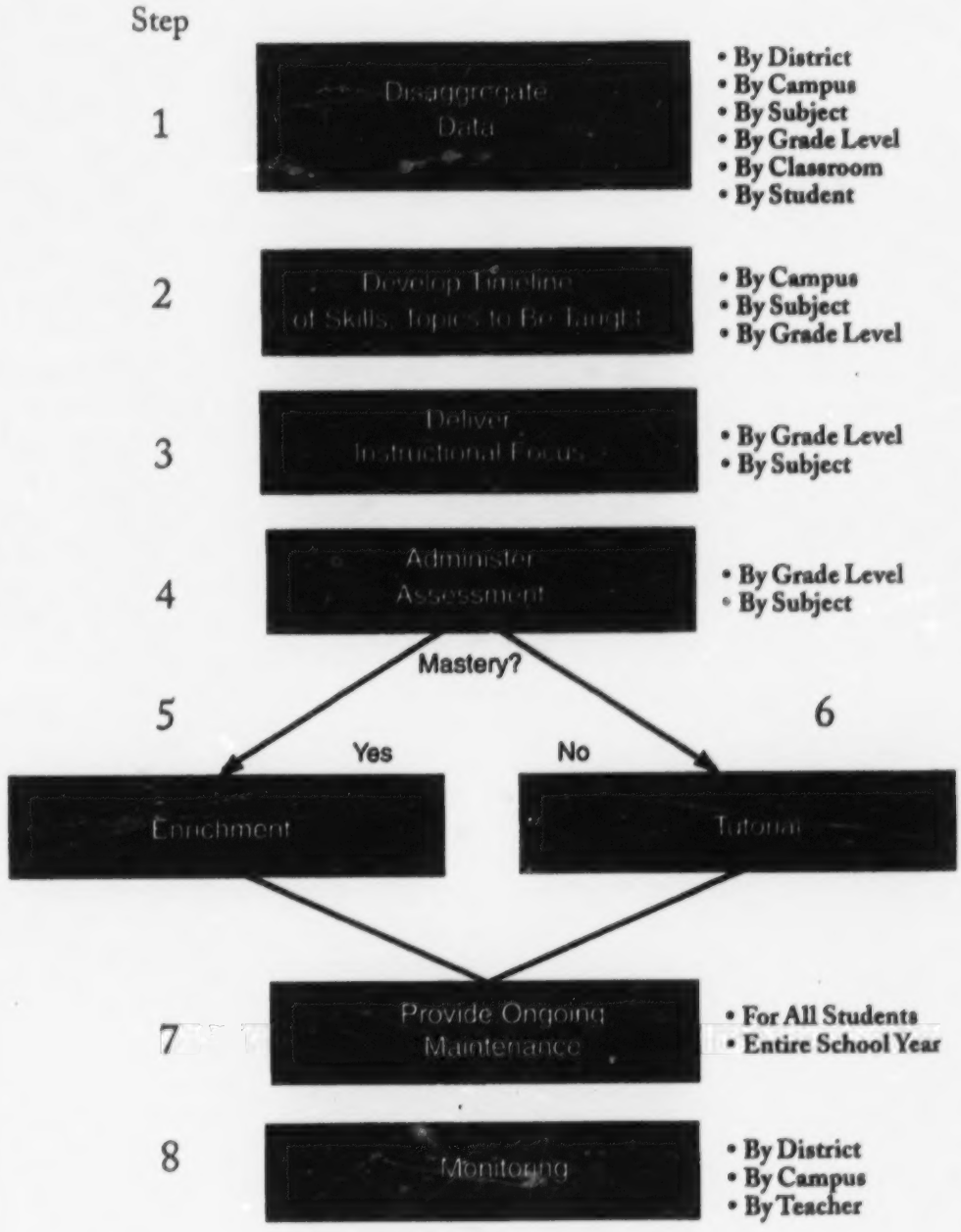
Today, teachers develop class-by-class academic calendars based on data provided both by TAAS and additional ongoing assessments, with some campuses having as many as nine of these teacher-developed benchmark tests. The calendars change depending on what information the assessments provide, with teachers expected to continually ask: "What do the students need to know? What do I need to teach? What are we going to do today? How much time will I need to do that?"

The data are also used to identify which students need extra instruction. To provide this, every school has developed an extensive support system, ranging from after-school tutorials to daily regrouping for skills-based instruction to computerized reading programs. Additional time for learning and practice comes during extended-day instruction after school, Saturday classes, an after-school Day Care program that provides tutoring, and an extended school year of 22 more days for those students who have not mastered the basic skills considered necessary at their grade level.

Students are also taught how to analyze and use their own assessment data. For example, profiles of the periodic math assessments used in one middle school are reviewed to see what specific objectives haven't yet been mastered. In many of the schools, even at the elementary level, students participate in "TAAS Talks" that help to pinpoint what still needs to be learned.

Source: Cawelti and Protheroe 2001.

Brazosport Independent School District
8-Step Instructional Process



BEST COPY AVAILABLE

- Which solution is most supported by the data about...
  - the challenge area;
  - why it exists;
  - potential outcomes that will help us move towards our schoolwide goals?
- How could we build support from parents, students, teachers, instructional assistants, community members, or other stakeholders for implementing this solution?
- Which teachers, departments, or grade levels are willing to implement this solution?
- How will we know if this solution is successful? Have we laid out a plan with clearly defined responsibilities for implementing the solution and assessing its effectiveness (i.e., who will do what by when...)?

Implement. The challenge for a school is to take action to improve student learning based on what the data reveal. Schools may change goals, instruction, curriculum, technical structures, or materials; reallocate resources; create new programs; or focus on a particular grade level and subject, such as fourth-grade literacy.

These actions must be directly related to the articulated goals of the school and the evidence gathered through the accountability process.

- Who is actually implementing the plan? Are they aware of the challenge being addressed? Are they helping to assess the results?
- How is data collected about the effectiveness of the plan? When should we expect results from the plan we have implemented? Is there a way to get a "dipstick" measurement of its effectiveness even if we have to wait longer for full results?
- What will we do if we get new information midstream—information that may impact the implementation of our plan?

Reflect and evaluate. Here the cycle begins anew. Participants realize that any actions or changes reflect their best guess—albeit a well-informed one—about what will bring about the desired improvements. So they find ways to analyze the effects of these changes. Questions asked include:

- What level of attainment will satisfy us that our changes have worked? What will constitute evidence?
- What does the data say about the success of the plan? Have we helped the school to move towards its goals for high achievement? Has data been disaggregated by ethnicity, grade level, etc. to determine the success of the

plan for all students?

- What do different stakeholders in the school think about the success of the plan? What have we learned from parents, students, teachers, instructional assistants, and community members about the success of the plan?
- What are our continued questions?
- What could we do to expand or improve the implementation?
- If the plan was unsuccessful, what does the data say about why the challenge still exists? What should we do next? (Annenberg Institute for School Reform undated; Coalition of Essential Schools 1996)

Program evaluation cannot be considered successful until results are used to improve instruction and student success. The context for continuous improvement implies that schools should never be satisfied with the programs. Improvements can always be made, and evaluation is an important tool for accomplishing this purpose (Fleischman and Williams 1996).

Notes, Reminders, and Ideas:

The Promise of Technology to Improve Use of Assessment Data

[We must]...directly link students' digital portfolio artifacts to the standards for which they demonstrate achievement. High technology disconnected from a focus on curriculum standards will only exacerbate the lack of meaningful integration of technology into teaching and learning.

—Barrett, undated

The promise of technology is the ability to effectively collect, store, make sense of, and report on various aspects of an educational program in real time. At any one moment that might encompass: a digital portfolio, a computerized gradebook, an electronic lesson planner, a set of district-approved assessments, a longitudinal database of student performance data, or a recent assessment with scores charted in such a way as to facilitate analysis.

The digital promise allows principals, teachers, and schools to focus on goals and to easily evaluate and report progress relative to these goals. It builds the capacity to support ongoing and reflective assessment practices. It allows the teacher and the school to determine what worked and what did not work. The technology provides the ability to maintain ongoing, multi-dimensional portfolios of what students are learning by managing the information in a way that is easily accessible.

We now have the technology tools to implement systems that let teachers and administrators look at performance over time and at several levels—individual student, classes, schoolwide, and districtwide. Such tools could provide access to proven resources from other educators—a living bank of proven, standards-referenced instructional materials, lessons, units, and assessments perfected through action research.

The digital promise allows an online dialogue where teachers and schools can share concerns and successes. This forum allows the mentoring process to happen anytime and any place. Interaction can take place in real time—"chat," or through an email exchange. Or, it can be as structured as a discussion in the "forum." For the first time, teachers truly have the opportunity to collaborate on lessons and units with a wide range of fellow practitioners.

The underlying question becomes, How shall we prepare our staff, as well as our curriculum, instruction, and assessment practices, to take full advantage of the efficiency of such systems, remembering that the key questions are:

- How is this student doing?
- How is the class doing?

How Technology-Savvy Are Your Teachers' Portfolios?

Have you given thought to your teachers' instructional portfolios recently? With advances in computer technology, there are new approaches to storing, organizing, and using portfolios. Consider the following questions as you reflect on your teachers' portfolios.

- How do you currently store assessment scores and artifacts? How accessible are audio, video, multimedia, and performance assessment materials?
- Do you have timely access to aggregated/disaggregated data for comparison purposes? How do you sort, organize, and display the data so that it facilitates analysis?
- How many times do you repeat the same data output—report card heading information, district and school forms, IEPs, lesson plans, or student study contracts?
- How can you streamline and organize anecdotal observations of student learning?
- How do you archive successful lesson plans, activities, and assessments? Is the majority of your professional library at home or at work?
- Do teachers, parents, and students have immediate access to many examples of student work and exemplary work samples?
- How do you communicate progress and/or assessment results to parents? To students?
- How do you manage the enormous amounts of data that schools are required to collect as well as the additional data the school wants to collect?

- Is accountability high on your priority list? Do you have your own copy of the standards for each subject area in your school? Do you reference them when planning, teaching, and learning?
- How do you manage the grading and reporting process?
- How and when do you collaborate with other educators? Look at student work? Share assessment practices and insights? Share professional dialogue?
- What's missing in your current method of storage and retrieval of data about student learning?

Notes, Reminders, and Ideas:

Conclusion

A key element characteristic of successful schools, in study after study, is the well-organized approach to using assessment data that these schools have developed. This did not happen overnight. Typically, it was an evolutionary process and one that may have included some false starts. Always, the principal played a role as instructional leader guiding the process.

A good starting place suggested by Hibbard and Yakimowski (1997) is spending some time developing the "guiding questions" for the effort:

1) What should students know and how should they be able to use what they know?
2) How well should students perform?
3) What will we do to assess student performance?
4) How well do students actually perform?
5) Based on our assessment results, what will we do to improve student performance?

In their view, these questions should lead teachers and other staff members to "purposeful conversations...about improving student performance" (1997, 67) that can use assessment as a resource.

All this takes time for school leaders and their staffs, but it *is* worth the effort. In its report, *Measuring What Matters*, the Committee for Economic Development puts its full support behind the use of assessment data to improve instruction:

> The evidence is mounting that measurement systems designed to enhance instruction, when accompanied by efforts to give teachers the capacity to use the results, can be powerful levers for

> instructional improvement.... Creative superintendents and principals use the information provided by assessments to spur teachers to identify the strengths and weaknesses of individual students and make plans for addressing problems. Teachers are stimulated to work together on new curricular models that align lesson plans to content standards and that encourage curricula to flow from subject to subject and from grade to grade within schools. Good teachers use test results to help them learn how better to teach the curriculum, not to teach to the test itself (2001, 15).

Effective use of assessment data to improve instruction is a complex task. To do it well, school staff need high-quality data, training and support in using the data, and time to collaborate. The desired end result is a system that integrates instruction and assessment in a cycle of continuous improvement.

Notes, Reminders, and Ideas:

Web Resources for Using Assessment Data

"A Framework for Accountability" (Annenberg Institute for School Reform)
www.annenberginstitute.org/accountability/framework/pgone.html

Achieve: Resource for the creation and revision of academic standards and assessments; includes a "clearinghouse" of standards.
www.achieve.org/

"Alternative Assessment Guidebook" (CRESST)
cresst96.cse.ucla.edu/CRESST/Sample/GBTHREE.PDF

"Alternative Assessment in K-12 Science Education" (Eisenhower National Clearinghouse)
www.enc.org/professional/research/journal/science/documents/0,1944,ACQ-103048-3048,00.shtm

"Alternative Assessment in the Social Sciences" (Illinois State Board of Education)
www.coe.ilstu.edu/jabraun/socialstudies/assess/socsci/

"Alternative Assessment Strategies" (RMC Research Corporation)
www.rmcdenver.com/useguide/assessme/definiti.htm

"An Introduction to Program Evaluation for Classroom Teachers." Howard L. Fleischman and Laura Williams. Development Associates, Inc.
www.teacherpathfinder.org/School/Assess/assess.html

Analysis of state standards
www.edexcellence.net/standards/best.html

Assessment Information Links for Teachers (Teacher/Pathfinder)
teacherpathfinder.org/School/Assess/assessmt.html

"Assessment Strategies and Definitions" (RMC Research Corporation)
www.rmcdenver.com/useguide/assessme/definiti.htm

Collection of standards in browseable and searchable formats
whyfiles.org/teach/4.html

CRESST: National Center for Research on Evaluation, Standards, and Student Testing
cresst96.cse.ucla.edu/

"Drawing Value from Evaluation"
www.ssta.sk.ca/research/evaluation_and_reporting/9d

Educational Testing Service, the United States' largest testing organization, provides useful analysis of test results with implications for education policy.
www.ets.org/research/pic/index.html

Eisenhower National Clearinghouse provides information on state and national standards with a focus on science and math.
www.enc.org/professional/standards/

"Ensuring Equity with Alternative Assessments" (North Central Regional Educational Laboratory)
ncrel.org/sdrs/areas/issues/methods/assment/as800.htm

ERIC Clearinghouse on Assessment and Evaluation
www.ericae.net/

ERIC Hotlist for Web-based Resources on Assessment
ericae.net/nintbod.htm

FairTest: The National Center for Fair & Open Testing
www.fairtest.org/

How Can I Ensure the Integrity of My Assessments? Guidelines for Appropriate Assessment (RMC Research Corporation)
www.rmcdenver.com/useguide/assessme/gideline.htm

Information on state standards (Fordham Foundation)
www.edexcellence.net/topics/standards.html

Kathy Schrock's information on alternative assessment; good information on rubrics
school.discovery.com/schrockguide/assess.html

Leadership Planning Sequence (Assessment Training Institute)
www.assessmentinst.com/

Links to state standards
www.education-world.com/standards/state/index.shtml

"Looking at Student Work" (project of the Annenberg Institute for School Reform)
www.lasw.org/

"Los Angeles Learning Center Alternative Assessment Guidebook" (National Center for Research on Evaluation, Standards, and Student Testing-CRESST)
cresst96.cse.ucla.edu/CRESST/Sample/GBTHREE.PDF

Mid-continent Research for Education and Learning; resources on standards (some for purchase; some available online)
www.mcrel.org/products/standards/index.asp

National Assessment of Educational Progress
nces.ed.gov/nationsreportcard/sitemap.asp

North Central Regional Educational Laboratory resource page on assessment
www.ncrel.org/sdrs/areas/as0cont.htm

NWREL Assessment Toolkit (Northwest Regional Educational Laboratory)
www.nwrel.org/assessment/toolkit98.asp

"Performance-Based Assessment for Accountability Purposes"
www.cresst96.cse.ucla.edu/CRESST/Reports/TECH390.PDF

"Portfolio Assessment and High Technology"
cresst96.cse.ucla.edu/CRESST/Sample/HIGHTECH.PDF

"Portfolio Trouble-Shooting" (RMC Research Corporation)
www.rmcdenver.com/useguide/assessme/portfoli.htm

Sample Performance Assessments for History, Science, and Mathematics
www.cse.ucla.edu/CRESST/pages/samples.htm

"What Does Research Say about Assessment?" (NCREL)
www.ncrel.org/sdrs/areas/stw_esys/4assess.htm

Recommended Books and Articles

Adams, T.L. 2000. "Alternative Assessment in Elementary School Mathematics." *Childhood Education* Vol. 74: 220-24.

Banta, T.W., and C.A. Palomba. 1999. *Assessment Essentials: Planning, Implementing, Improving.* San Francisco, CA: Jossey-Bass Publishers.

Burke, K. 1999. *How to Assess Authentic Learning*, 3rd. edition. Englewood Cliffs, NJ: Prentice Hall.

Calfee, R., and M. Gearhart. 1998. "Introduction: Portfolios and Large-Scale Assessment." *Educational Assessment* Vol. 5: 1-3.

Chen, Y., and M.A. Martin. 2000. "Using Performance Assessment and Portfolio Assessment Together in the Elementary Classroom." *Reading Improvement* Vol. 37: 32-38.

Daniels, V. I. 2000. "The Assessment Maze: Making Instructional Decisions about Alternative Assessments for Students with Disabilities." *Preventing School Failure* Vol. 43: 171-78.

Darling-Hammond, L., and J. Snyder. 2000. "Authentic Assessment of Teaching in Context." *Teaching and Teacher Education* Vol. 16: 523-45.

Fourie, I., and D. van Niekerk. 1999. "Using Portfolio Assessment in a Module in Research Information Skills." *Education for Information* Vol. 17: 333-352.

Gearhart, M., and J.L. Herman. 1998. "Portfolio Assessment: Whose Work Is It? Issues in the Use of Classroom Assignments for Accountability." *Educational Assessment* Vol. 5: 41-55.

Gronlund, N.E., R.L. Linn, and K.M. Davis. 1990. *Measurement and Assessment in Teaching.* Englewood Cliffs, NJ: Prentice Hall.

Herman, J. 1997. "Assessing New Assessments: How Do They Measure Up?" *Theory into Practice* (Autumn 1997).

Herman, J. L., P.R. Aschbacker, and L. Winters. 1992 . *A Practical Guide to Alternative Assessment.* Alexandria, VA: Association of Supervision & Curriculum Development.

Kagan, S.L. 2000. "Making Assessment Count...What Matters?" *Young Children* Vol. 55: 4-5.

National Forum on Assessment. 1995. *Principles and Indicators for Student Assessment Systems.* Cambridge, MA: FairTest.

Siegel-Causey, E., and R.M. Allinder. 1998. "Using Alternative Assessment for Students with Severe Disabilities: Alignment with Best Practices." *Education and Training in Mental Retardation and Developmental Disabilities* Vol. 33: 168-78.

Staytor, F., and P. Johnston. 1990. "Evaluating the Teaching and Learning of Literacy." In *Reading and Writing Together: New Perspectives for the Classroom*, Timothy Shanahan, editor. Norwood, MA: Christopher-Gordon Publishers.

Stiggins, R.J. 2000. *Student-involved Classroom Assessment.* Englewood Cliffs, NJ: Prentice Hall.

Stiggins, R.J., and N.F. Conklin. 1992. *In Teachers' Hands: Investigating the Practices of Classroom Assessment.* Albany: State University of New York Press.

Tombari, M., and G. Borich. 1999. *Authentic Assessment in the Classroom: Applications and Practice.* Englewood Cliffs, NJ: Prentice Hall.

Wiggins, G. 1993. *Assessing Student Performance: Exploring the Purpose and Limits of Testing.* San Francisco: Jossey-Bass Publishers.

Wiggins, G. 1998. *Educative Assessment: Assessment to Improve and Inform Learning.* San Francisco: Jossey-Bass Publishers.

Wiggins, G. 1997. "Work Standards: Why We Need Standards for Instructional and Assessment Design." *NASSP Bulletin* Vol. 81, No. 590: 56-64.

An Assessment Glossary

Accountability
The system of structures, policies, and practices put into place to meet the demand by a community (public officials, employers, and taxpayers) for school officials to prove that money invested in education has led to measurable learning.

Accountability testing is an attempt to sample what students have learned, or how well teachers have taught, and/or the effectiveness of a school's or principal's performance as an instructional leader.

Achievement Target
An educational goal to be attained through instruction, hence the prime focus of any assessment.

Achievement Test
A standardized test designed to measure the things a student knows and can do in such subjects as reading, spelling, or mathematics, usually as a result of classroom instruction. Such testing produces a statistical profile used as a measurement to evaluate student learning in comparison with a standard or norm.

Action Plans
Plans that refer to principal school-level actions, derived from short- and long-term strategic planning. In simplest terms, action plans spell out those things the school must do well for its strategy to succeed.

Action Research
School and classroom-based studies initiated and conducted by teachers and other school staff. Action research involves teachers, aides, principals, and other school staff as researchers who systematically reflect on their teaching or other work and collect data that will answer their questions. It offers staff an opportunity to explore issues of interest to them in an effort to improve classroom instruction and educational effectiveness.

Age Norms
Values representing typical or average performance of individuals in a specific age group.

Alternative Assessment
See *Authentic Assessment.*

Analytic Trait
The use of separate criteria in scoring work, typically involving separate rubrics for each key criterion; also called primary-trait scoring. In effect, a performance is assessed numerous times, using the "lens" of a separate criterion each time. Analytic-trait scoring is in contrast with holistic scoring whereby a judge forms an overall impression about a performance.

Anchors
The samples of work or performance used to set the specific performance standard for each level of a rubric. Thus, attached to the paragraph describing a 6-level performance in writing would be 2-3 samples of writing that illustrate what a 6-level performance looks like. (The anchor for the top score is often called the "exemplar.")

Aptitude Test
A test intended to measure the test-taker's innate ability to learn, given before receiving instruction.

Assess, Assessment
In an educational context, the process of observing learning—describing, collecting, recording, scoring, and interpreting information about a student's or one's own learning. Assessment techniques include tests, exhibits, interviews, surveys, observation, etc.

Assessment Literacy
The possession of knowledge about the basic principles of sound assessment practice, including terminology, the development and use of assessment methodologies and techniques and familiarity with standards of quality in assessment. Increasingly, this also includes familiarity with alternatives to traditional measurements of learning.

Authentic Assessment
An authentic assessment is composed of tasks and activities designed to simulate or replicate important, real-world challenges. The heart of authentic assessment is performance-based testing—asking the student to use knowledge in a realistic way. Thus, the context of the assessment, not just the tasks ("messiness" of problem, ability to seek feedback and revise, access to apt resources, etc.), must be more realistic than conventional testing. Authentic assessment can take place at any point in the learning process.

Average
A statistic that indicates the central tendency or most typical score of a group of scores. Most often average refers to the sum of a set of scores divided by the number of scores in the set.

Benchmark
A measurement of performance against an established standard at defined points along the path toward the standard. In many districtwide systems, there are "benchmarks" set for grades 4, 8, 10, and 12, for example.

A benchmark can also set the highest possible standard of performance. Thus, a benchmark in this sense is student achievement that illustrates points on a performance scale, used as exemplars.

Bias
The lack of objectivity, fairness, or impartiality in assessment. A judge of student performance is technically biased if a) his/her standards are higher or lower than the agreed-upon ones, b) he/she tends to focus on particular strengths or weaknesses in the performance that are not consistent with specified guidelines, or c) he/she focuses on the performer's personal traits or characteristics vs. the qualities of the performance or product to be judged.

Bloom's Taxonomy
A scheme developed by Benjamin Bloom for distinguishing the simplest forms of recall from the most sophisticated uses of knowledge in designing student assessments. The six elements were called Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation.

Ceiling
The upper limit of ability that can be measured by a particular test.

Cohort
A group whose progress is followed by means of measurements over time.

Competency Test
A test intended to establish that a student has met established minimum standards of skills and knowledge and is thus eligible for promotion, graduation, certification, or other official acknowledgment of achievement.

Constructed-Response Questions
A type of performance assessment problem that requires students to produce their own answers rather than select from an array of possible answers. Such questions may have just one correct response, or they may be more open-ended, allowing a range of responses.

Content Standards
Specifications of the general domains of knowledge that students should learn in various subjects.

Criteria
The qualities that must be met by performances/products for work to be up to standard and for the performances to be deemed successful. The assessment must also determine how much weight each criterion should receive relative to other criteria in making a judgment.

Criterion-Referenced Test
An approach that focuses on whether a student's performance meets a predetermined standard level, usually reflecting mastery of the skills being tested. The focus is on performance of an individual as measured against a standard or criteria rather than against performance of others who take the same test, as with norm-referenced tests.

Curriculum Alignment
The degree to which a curriculum's scope and sequence matches a testing program's evaluation measures, thus ensuring that teachers will use successful completion of the test as a goal of classroom instruction.

Curriculum-Embedded or Learning-Embedded Assessment
Assessment that occurs simultaneously with learning, such as projects, portfolios and "exhibitions," and occurs in the classroom setting. If properly designed, students should not be able to tell whether they are being taught or assessed. Tasks or tests are developed from the curriculum or instructional materials.

Descriptor
A set of signs used as a scale against which a performance or product is placed in an evaluation. Descriptors allow assessment to include clear guidelines for what is and is not valued in student work.

Diagnostic Test
An intensive, in-depth evaluation process with a relatively detailed and narrow coverage of a specific area. The purpose of this test is to determine the specific learning needs of individual students and to be able to meet those needs through regular or remedial classroom instruction.

Dimensions, Traits, or Subscales
The subcategories used in evaluating a performance or portfolio product (e.g., in evaluating student writing, one might rate student performance on subscales such as organization, quality of content, mechanics, or style). Separate descriptors or scoring methods may apply to each dimension of the student's performance assessment.

Disaggregated Data
Data that has been broken out to illuminate the performance of different groups of students. For example, disaggregated data may show how one population performs, such as how boys perform, how language minority students perform, how African American students perform, etc.

Essays
A type of performance assessment used to evaluate a student's understanding of the subject through a written description, analysis, explanation, or summary. It can also be used to evaluate a student's composition skills. Responses can be brief or extensive.

Evaluation
Using collected information (assessments) to make informed decisions about continued instruction, programs, and activities.

Formative Assessment
Observations that allow one to determine the degree to which students know or are able to do a given learning task, and that identify the part of the task that the student does not know or is unable to do. Outcomes suggest future steps for teaching and learning.

Grade Equivalent
A score that describes student performance in terms of the statistical performance of an average student at a given grade level.

Grading
Allocating students to an ordered set of named categories ordered by merit; also known as rating. Grades are generally criterion-referenced.

Higher-Order Skills
Abilities used for more sophisticated tasks requiring application of knowledge to solve a problem.

High-Stakes Testing
Any testing program whose results have important consequences for students, teachers, schools, and/or districts. Such stakes may include promotion, certification, graduation, or denial/approval of services and opportunity.

Holistic Scoring
A scoring system that is the result of an overall impression of the quality of a performance or portfolio product. The rater matches his or her impression with a point scale, generally focusing on specific aspects of the performance or product. Holistic scoring is distinguished from analytic trait scoring, where separate rubrics are used for each separate criterion that makes up an aspect of performance.

Ill-Structured Question or Task
A question for which there is no recipe or obvious formula to answer it; or for which no specific strategy for success is stated or implied. Such questions thus demand more than knowledge: they demand good judgment and imagination. All good essay questions, science problems, or design challenges are thus ill-structured: even when you understand the goal or know what is expected, you have to scratch your head and build not merely an answer but a procedure as you go.

Informal Test
A non-standardized test that is designed to give an approximate index of an individual's level of ability or learning style; often teacher-constructed.

Instructional Assessment
Evaluation of class progress relevant to the curriculum.

Instrument
A measuring device (e.g. test) used to determine the present value of something under observation.

Inventory
A catalog or list for assessing the absence or presence of certain attitudes, interests, behaviors, or other items regarded as relevant to a given purpose.

I.Q. (Intelligence Quotient) Test
The first of the standardized norm-referenced tests, developed during the nineteenth century. Traditional psychologists believe that neurological and genetic factors underlie "intelligence" and that scoring the performance of certain intellectual tasks can provide assessors with a measurement of general intelligence. There is a substantial body of research that suggests that I.Q. tests measure only certain analytical skills, missing many areas of human endeavor considered to be intelligent behavior. I.Q. is considered by some to be fixed or static; whereas an increasing number of researchers are finding that intelligence is an ongoing process that continues to change throughout life.

Item Analysis
Analyzing each item on a test to determine the proportions of students selecting each answer. Can be used to evaluate student strengths and weaknesses; may point to problems with the test's validity and to possible bias.

Journals
Students' personal records and reactions to various aspects of learning and developing ideas. A reflective process often found to consolidate and enhance learning.

Longitudinal Assessment
Also known as "developmental" assessment; involves assessing the same performances over numerous times, using a fixed scoring continuum, to track progress (or lack of it) toward a standard. For example, the National Assessment of Educational Progress (NAEP) uses a fixed scale for measuring gains in mathematics performance over the 4th, 8th, and 12th grade. Most school testing, whether done locally or statewide, is not longitudinal since the tests are one-time events with one-time scoring systems.

Lower-Order Skills
Abilities used for less sophisticated tasks, such as recognition, recall, or simple deductive reasoning.

Mastery Level
The level of performance actually needed on a criterion; sometimes the level thought to be optimal and feasible.

Matrix Sampling
A strategy used in large-scale testing programs to get better information about overall performance by using a variety of tests during the same test administration. For example, in a writing test, seven different genres of writing might be identified. In the test, each student might get only one writing task focused on one genre. But across all students tested, all genres would be tested.

Mean
The average. Add all the scores and divide that sum by the number of people taking the test to find the mean score of a test.

Measurement
Determination of the magnitude of a quantity; includes numerical scoring.

Mean
One of several ways of representing a group with a single, typical score. It is figured by adding up all the individual scores in a group and dividing them by the number of people in the group. The mean can be affected by extremely low or high scores.

Median
One of several ways to represent a group's scores with a single, typical score. The median is the point on a scale that divides a group into two equal subgroups; the middle raw data point; the "middle" performance. The median is not affected by low or high scores as is the mean.

Metacognition
The knowledge of one's own thinking processes and strategies, and the ability to

consciously reflect and act on the knowledge of cognition to modify those processes and strategies.

Mode
The most frequent score or score interval; there can be more than one mode.

Mode, Performance Mode
Refers to the type or method of performance being demonstrated and assessed. Modes of performance include written, oral, and visual. Within each mode there are many genres.

Multidimensional Assessment
Assessment that gathers information about a broad spectrum of abilities and skills as in Howard Gardner's theory of Multiple Intelligences.

Multiple-Choice Test (or Item)
Test or test item that gives students the opportunity to select responses to test questions from among a number of specific choices. This format, called "selected-response" or "multiple-choice," is efficient and practical. Carefully designed multiple-choice questions can provide valid information about students' knowledge and their ability to reason logically and apply complex thinking processes to solve problems. Norm-referenced tests are usually administered in a multiple-choice format, where the correct answer is provided along with incorrect answers. These are the tests most adults remember taking in their youth. In most instances, multiple-choice tests are scored by computers and provide impartial, accurate results.

Norm
Performance standard that is established by a reference group and that describes average or typical performance. The norm is the midpoint (or median) of scores or performance of the students in that group. Fifty percent will score above and fifty percent below the norm.

Normal Curve Equivalent
A score that ranges from 1-99, often used by testers to manipulate data arithmetically. Used to compare different tests for the same student or group of students and between different students on the same test. An NCE is a normalized test score with a mean of 50 and a standard deviation of 21.06. NCEs should be used instead of percentiles for comparative purposes.

Norm Group
A random group of students selected by a test developer to take a test to provide a range of scores and establish the percentiles of performance for use in establishing scoring standards.

Norming Population
The group of students to whom the test results are being compared to create percentile numbers. For some tests, the norming population will be all students at the same grade level across the state or country. For other tests, the norming population will only be students with a similar demographic.

Norm-Referenced Tests
Kind of tests that measure basic concepts and skills commonly taught in schools throughout the country. These tests are not designed as precise measures of any given curriculum or single instructional program. Results from norm-referenced tests provide information that compares students' achievement with that of a representative national sample. This gives teachers the opportunity to compare their students with other students. So, when a teacher says that a student scored at the 82nd percentile, that student's score was equal to or better than 81 percent of the scores of all the students who took the same norm-referenced test during the norming process. This approach does not focus on absolute levels of mastery, as with criterion-referenced approaches.

Objective Percent Correct
The percent of the items measuring a single objective that a student answers correctly.

Objective Test
A test for which the scoring procedure is completely specified enabling agreement among different scorers; a correct-answer test.

On-Demand Assessment
An assessment process that takes place as a scheduled event outside the normal routine; an attempt to summarize what students have learned that is not embedded in classroom activity.

Open-Ended Problems, Tasks, or Questions
Problems, tasks, or questions that do not lead to a single "right" answer. This does not imply that all answers are of equal value, however. Rather, it implies that many different acceptable answers are possible. Such answers are thus "justified" or "plausible" or "well-defended" as opposed to "correct." Essay test questions, for example, are all open-ended. By contrast, all multiple-choice tests are not open-ended by design.

Opportunity-to-Learn Standards
Criteria for evaluating whether schools are giving students the chance to learn material reflected in the content standards. This may include such specifics as the availability of instructional materials or the preparation of teachers.

Oral Presentations/Interviews
Type of performance assessment that allow students to verbalize their knowledge.

Outcome
Used in education as shorthand for "intended outcomes of instruction." An "intended outcome" is a desired result, specific goal or performance that can be measured. Stiggins uses the term "achievement target" to describe such intents.

Percentile
A rank order based on a scale of 100. A percentile rank indicates the percentage of a reference or norm group obtaining scores equal to or less than the test-taker's score. For example, a student scoring at the 70th percentile scored higher than 69% of the students who took the test. A percentile score does not refer to the percentage of questions answered correctly, it indicates the test-taker's standing relative to the norm group standard.

Differences in percentile points are not equal throughout the scale (e.g., the difference between 45th and 50th percentile is not equal to the difference between 1st and 5th percentile). As a result, percentiles cannot be averaged, summed, or combined in any way.

Percentiles are based on 100 divisions or groupings, deciles are based on 10, and quartiles are based on 4 groups.

Percent Score
The percent of items that are answered correctly.

Perform (Performance)
To use one's knowledge to effectively act or bring to fruition a complex product in which one's knowledge and expertise is revealed. Music recitals and auto mechanic competitions are performances in both senses; so are oral exams.

Performance Assessment
An evaluation in which students are asked to engage in a complex task, often involving the creation of a product. Student performance is rated based on the process the student engages in and/or based on the product of his/her task. Many performance assessments emulate actual workplace activities or real-life skill applications that require higher order processing skills. Performance assessments can be individual or group-oriented.

On a performance assessment, students may be asked to write an essay or short response, draw a conclusion, respond to a reading passage, or perform a science experiment. Many performance assessments call on students to produce solutions to "constructed response" (short-answer) items or open-ended questions. Teachers or other school personnel observe students' performances and rate the outcomes. This

kind of assessment is also useful in measuring listening skills, writing, and the process of problem solving. Performance assessments can also be standardized so that the test is given and scored the same way at each administration.

Performance-Based Assessment
Direct observation and rating of a student performance of an educational objective; may be over a period of time, and typically involving the creation of a product.

Stiggins (1997) defines performance-based assessment as the use of performance criteria to determine the degree to which a student has met an achievement target.

Assessment of the performance is done using a rubric, or analytic scoring guide to aid in objectivity. Performance-based assessment is a test of the ability to apply knowledge in a real-life setting.

The assessment may be a continuing interaction between teacher and student and should ideally be part of the learning process. *See Performance Assessment.*

Performance Criteria
A predetermined list of observable standards used to rate performance assessments. Effective performance criteria include considerations for validity and reliability and provide students with important information about expectations, giving them a target or goal to strive for.

Performance Standards
External criteria used to establish the degree or quality of students' performance in the subject area set out by the content standards, answering the question "How good is good enough?"

Portfolio
A collection of representative student work that exhibits to others the direct evidence of a student's efforts, achievements, and progress over a period of time.

Generally, the collection should involve the student in selection of its contents, and should include information about the performance criteria, the rubric or criteria for judging merit, and evidence of student self-reflection or evaluation. It should include representative work, providing a documentation of the learner's performance and a basis for evaluation of the student's progress. Portfolios may include a variety of demonstrations of learning and have been gathered in the form of a physical collection of materials, videos, cd-roms, reflective journals, etc.

In academic subject areas such as English/Language Arts or Mathematics, a portfolio often serves two distinct purposes: providing a documentation of the student's work, and serving as the basis for evaluation of work-in-progress or work over time. The

documentation typically serves three functions: revealing the student's control over all the major areas/techniques/genres/topics of the course or program, allowing students to reflect on and show off their best work (by letting them select which works will be put in the portfolio), and providing evidence of how works evolved and were refined.

Portfolios may be used for evaluation of a student's abilities and improvement.

Portfolio Assessment
Type of performance assessment that involves the ongoing evaluation of a cumulative collection of creative student works. Portfolios may be assessed in a variety of ways. Each piece may be individually scored, or the portfolio might be assessed merely for the presence of required pieces, or a holistic scoring process might be used. It can include student self-reflection and monitoring.

It is common that assessors work together to establish consensus of standards to ensure greater reliability in evaluation of student work. Established criteria are often used by reviewers and students involved in the process of evaluating progress and achievement of objectives.

Practice Effect
The improved performance produced by taking a second test with the same or closely similar items, even if no additional learning has occurred between the tests.

Primary Trait Method
A type of rubric scoring constructed to assess a specific trait, skill, behavior, or format, or the evaluation of the primary impact of a learning process on a designated audience.

Process
In the context of assessment, process refers to the intermediate steps the student takes in reaching the final performance or end-product specified by the assessment. "Process" thus includes all strategies, decisions, sub-skills, rough drafts, and rehearsals used in completing the given task.

Process can be evaluated as part of an assessment, as in the example of evaluating a student's performance during pre-writing exercises leading up to the final production of an essay or paper.

Product
The tangible and stable result of a performance or task. An assessment is made of student performance based on evaluation of the product of a demonstration of learning.

The product is valid for assessing the student's knowledge to the extent that success or failure in producing the product a) is dependent upon the knowledge we taught and

want to assess, and b) appropriately "samples" from the whole curriculum in a way that mirrors the relative importance of the material in the course.

Profile
A graphic compilation of the performance of an individual on a series of assessments.

Program Evaluation
Assessment of the effectiveness of particular instructional interventions or programs.

Project
A complex assignment involving more than one type of activity and production. Projects can take a variety of forms; some examples are a mural construction, a shared service project, or other collaborative or individual effort.

Prompt
An assignment or directions asking the student to undertake a task or series of tasks. A prompt presents the context of the situation, the problem or problems to be solved, and criteria or standards by which students will be evaluated.

Published Test
A test that is publicly available because it has been copyrighted and published commercially.

Qualitative Data
Data of a generally descriptive nature, which is difficult to quantify. Methods for gathering qualitative data include observations, case studies, ethnography, collecting student work samples, open-ended surveys, and interviewing.

Quantitative Data
A term that generally refers to numerical data. Common forms of quantitative data in schools are standardized test scores, multiple-choice surveys and student demographics.

Quartile
The breakdown of an aggregate of percentile rankings into four categories: the 0-25th percentile, 26-50th percentile, etc.

Quintile
The breakdown of an aggregate of percentile rankings into five categories: the 0-20th percentile, 21-40th percentile, etc.

Ranking
Placing students in an order, usually of merit, on the basis of their relative performance on a test, measurement, or observation.

Rating Scale
A written list of performance criteria related to a specific activity or product that an observer uses to assess student performance on each criterion in terms of its quality. The scale is usually based on descriptive words or phrases that indicate the performance levels "advanced, intermediate, and novice," in order to designate a level of achievement. The scale may be used with rubrics or descriptions of each level of performance.

Raw Score
The number of questions a student scored correctly on a given test, before being converted (e.g. to a percentile or grade equivalent).

Reliability
The extent to which an assessment is dependable, stable, and consistent when administered to the same individuals on different occasions or yields similar results over time with similar populations in similar circumstances. Two elements are looked at. The first seeks consistency in results. For example, will a student's score on a test today be close to his score tomorrow? The second, generalizability, seeks to ensure that assessment questions cover a subset of skills that can capture or "generalize" a broader universe of skills.

Rubric
A set of guidelines for giving scores. A typical rubric states all the dimensions being assessed, contains a scale, and helps the rater place the given work properly on the scale.

Sampling
A way to obtain information about a large group by examining a smaller, randomly chosen selection (the sample) of group members. If the sampling is conducted correctly, the results will be representative of the group as a whole. Sampling may also refer to the choice of smaller tasks or processes that will be valid for making inferences about the student's performance in a larger domain. "Matrix sampling" asks different groups to take small segments of a test; the results will reflect the ability of the larger group on a complete range of tasks.

Scale
A classification tool or counting system designed to indicate and measure the degree to which an event or behavior has occurred.

Scale Scores
Scores based on a scale ranging from 001 to 999. Scale scores are useful in comparing performance in one subject area across classes, schools, districts, and other large populations, especially in monitoring change over time.

Scaled Score
The raw score converted to a score that gives more weight to hard questions and less weight to easier questions.

Score
A rating of performance based on a scale or classification.

Scoring
1) A package of guidelines intended for people scoring performance assessments. May include instructions for raters, notes on training raters, rating scales, samples of student work exemplifying various levels of performance. Or, 2) Use of a numerical grade to rank student performance and assessment.

Scoring Criteria
Rules for assigning a score or the dimensions of proficiency in performance used to describe a student's response to a task. May include rating scales, checklists, answer keys, and other scoring tools.

Screening
A fast, efficient measurement for a large population to identify individuals who may deviate in a specified area, such as the incidence of maladjustment or readiness for academic work.

Self-Assessment
A process in which a student engages in a systematic review of a performance, usually for the purpose of improving future performance; may involve comparison with a standard, established criteria. May involve critiquing one's own work or may be a simple description of the performance. Reflection, self-evaluation, metacognition, are related terms.

Senior Project
Extensive projects planned and carried out during the senior year of high school as the culmination of the secondary school experience. Senior projects require higher-level thinking skills, problem-solving, and creative thinking. They are often interdisciplinary, and may require extensive research. Projects culminate in a presentation of the project to a panel of people, usually faculty and community mentors, sometimes students, who evaluate the student's work at the end of the year.

Standard
The performance level associated with a particular rating or grade on a given criterion or dimension of achievements. A performance standard is a specific result or level of achievement that is deemed exemplary or appropriate. Sometimes the word is used in education as a synonym for "high expectations." At other times, "standard" is used as a synonym for

"benchmark"—the best the performance or product can be done or has been done. And in large-scale testing, "standard" has often implicitly meant "minimal standard."

Standard Deviation
Statistical measure that tells the examiner the percentage of scores falling between points on a test's distribution of scores.

Standard Scores
Scores related to percentiles; a score that is expressed as a deviation from a population mean.

Standardized Achievement Tests
Type of test commonly used to provide valid, reliable, and unbiased information about students' knowledge in various areas. "Standardized" means that the test is always given and scored the same way. The same questions are asked and the same directions are given for each test. Specific time limits are set, and each student's performance may be compared with that of all the other students taking the same test. Most standardized achievement tests are norm-referenced, multiple-choice tests.

Standards
Agreed-upon values used to measure the quality of student performance, instructional methods, curriculum, etc.

Stanine
Shows a student's range of performance on a scale from 1-9. Scores of 1-3 are generally below average, scores of 4-6 are average, and 7-9 are above average.

Subjective Test
A test in which the impression or opinion of the assessor determines the score or evaluation of performance. The answers cannot be known or prescribed in advance.

Summative Assessment
Evaluation at the culmination of an activity, a plan, or a unit or subject of instruction to determine or judge student skills and knowledge or effectiveness of plan or activity.

Task
A goal-directed assessment activity, demanding that the students use their background of knowledge and skill in a continuous way to solve a complex problem or question. It typically involves diverse activities spread out over a lengthy period of time: doing research and bringing it to fruition in reports, building a museum exhibit, etc.

Teacher Assessment
Evaluation of a teacher's effectiveness conveying material, ideas, and new systems of thought to students. It requires evidence about the quality of what is taught, the amount that is learned, and the professionalism and ethicality of the teaching process.

Test
A measure that provides information about a person's knowledge, skill, competence, or behavior. Tests, or "assessments," are often used to monitor education systems for public accountability, help improve instructional practices, evaluate program effectiveness, measure student achievement, and assess student mastery of skills.

Test-Driven Curriculum
What results when teachers begin to teach to the test in order to prepare students for the content of the test, particularly when there are high stakes attached to the results of the test. This is not necessarily viewed as a problem if the assessment measures the desired student skills.

Testing
Measurement, in many cases any specific and explicit effort at performance or attitude evaluation, usually of students.

Validity
The extent to which an assessment measures what it was intended to measure.

References

Achieve, Inc. 1999. *Better Standards, Better Schools.* Online: www.achieve.org/achieve/achievestart.nsf/pages/annual

Airasian, P. 1994. *Classroom Assessment.* 2nd edition. New York: McGraw Hill Publishers.

American Association for the Advancement of Science. Undated. *Blueprints for Reform Online. Project 2061 of the American Association for the Advancement of Science.* Online: www.project2061.org/tools/bluepol/blpframe.htm

Anderson, S.R. 1993. "Trouble with Testing." *The American School Board Journal* Vol. 180, No. 6: 24-26.

Annenberg Institute for School Reform. Undated. *Toolbox for Accountability.* Online: www.annenberginstitute.org/accountability/toolbox/index.html

Archbald, D., and F. Newmann. 1988. *Beyond Standardized Testing: Authentic Academic Achievement in the Secondary School.* Reston, VA: NASSP Publications.

Armstrong, T. 1994. *Multiple Intelligences in the Classroom.* Alexandria, VA: Association for Supervision and Curriculum Development.

Arter, J.A., and V. Spandel. 1992. "NCME Instructional Module: Using Portfolios of Student Work in Instruction and Assessment." *Educational Measurement* Vol. 11, No. 1: 36-44.

Aschbacher, P.R., G. Koency, and J. Schacter. 1995. Los Angeles Learning Center Alternative Assessment Guidebook. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Online: cresst96.cse.ucla.edu/CRESST/Sample/GBTHREE.PDF

Assessment Training Institute. 2001. *Leadership for Excellence in Assessment: A School District Planning Guide.* Portland, Oregon: Author. Online: www.assessmentinst.com/pdfs/LeadershipGuide.pdf

Austin, T. 1994. *Changing the View: Student-led Parent Conferences.* Portsmouth, NH: Heinemann.

Baker, E.L., R.L. Linn, and J.L. Herman. 1996. *CRESST: A Continuing Mission to Improve Educational Assessment. Evaluation Comment.* Los Angeles: UCLA Center for the Study of Evaluation.

Bamburg, J.D. 1994. *NCREL Monograph: Raising Expectations to Improve Student Learning.* North Central Regional Educational Laboratory. Online: www.ncrel.org/sdrs/areas/issues/educatrs/leadrshp/le0bam.htm

Barber, J., L. Bergman, J. Goodman, K. Hosoume, L. Lipner, C. Sneider, and L. Tucker. 1995. *Insights & Outcomes: Assessments for Great Explorations in Math and Science.* Berkeley, CA: Lawrence

Hall of Science, Regents of the University of California.

Baron, J.B., and D.P. Wolf, editors. 1996. *Performance-Based Student Assessment: Challenges and Possibilities.* Chicago, IL: University of Chicago Press.

Barrett, H.C. Undated. *Electronic Portfolios and Standards.* Online: transition.alaska.edu/www/portfolios/TelEd98Abstract.html

Belanoff, P., and M. Dickson, editors. 1991. *Portfolios: Process and Product.* Portsmouth, NH: Boynton/Cook.

Birenbaum, M., and F. Douchy. 1996. *Alternative in Assessment of Achievements, Learning Processes and Prior Knowledge.* Online: www.nwrel.org/assessment/toolkit98/rubber.html

Black, P., and D. Wiliam. 1998. "Inside the Black Box: Raising Standards Through Classroom Assessment." *Phi Delta Kappan* (October 1998). Online: www.pdkintl.org/kappan/kbla9810.htm

Bloom, B.S., editor. 1954/1956. *Taxonomy of Educational Objectives. Book 1, Cognitive Domain.* New York: Longman, Inc.

Bloom, B.S., G.F. Madaus, and J.T. Hastings. 1981. *Evaluation to Improve Learning.* New York: McGraw Hill.

Bond, L., L. Friedman, and A. van der Ploeg. 1993. *Surveying the Landscape of State Educational Assessment Programs.* Washington, DC: Council for Educational Development and Research and the National Education Association.

Brandt, R.S. 1992. "Using Performance Assessment." *Educational Leadership*, 49: 8.

Brandt, R.S., editor. 1992. *Performance Assessment: Readings from Educational Leadership.* Alexandria, VA: Association for Supervision and Curriculum Development.

Burke, K., editor. 1992. *Authentic Assessment: A Collection.* Palatine, IL: IRI/Skylight Publishing.

California Assessment Program. 1989. *Guidelines for the Mathematics Portfolio: Phase II Pilot Working Paper.* Sacramento, CA: California Assessment Program Office (CAP), California State Department of Education.

California State Department of Education. 1989. *A Question of Thinking: A First Look at Students' Performance on Open-Ended Questions in Mathematics.* Sacramento, CA: California State Department of Education.

Cawelti, G., and N. Protheroe. 2001. *High Student Achievement: How Six School Districts Changed into High-Performance Systems.* Arlington, VA: Educational Research Service.

Center on Learning, Assessment, and School Structure. 1993. *Standards, Not Standardization. Vol. III, Rethinking Student Assessment.* Geneseo, NY: Author.

Coalition of Essential Schools. 1996. "The Tuning Protocol: A Process for Reflection on Teacher and Student Work." *Horace* (November 1996). Online: www.essentialschools.org/fieldbook/schdesign/inquiry/examstudwork/caprotocol.html

Colorado Department of Education. *Bringing Standards to Life: A "How-to" Guide to Contextual Learning and Curriculum Integration.* Online: www.aclin.org/sarb/linages/pdf/wkbkfnl.

Committee for Economic Development, Research and Policy Committee. 2001. *Measuring What Matters: Using Assessment and Accountability to Improve Student Learning.* New York: Committee for Economic Development. Online: www.ced.org/docs/education.pdf

Cromey, A., and M. Hanson. 2000. *An Exploratory Analysis of School-based Student Assessment Systems.* Naperville, IL: North Central Regional Education Laboratory. Online: www.ncrel.org/re/eval/mi/analysis.htm.

Cross, C.T. 1998. "The Standards Wars: Some Lessons Learned." *Education Week* (October 21, 1998). Online: www.edweek.org/ew/vol-18/08cross.h18

Darling-Hammond, L., J. Ancess, and B. Falk. 1995. *Authentic Assessment in Action: Studies of Schools and Students at Work*. New York: Teachers College Press.

Driscoll, M., and D. Bryant. 1998. *Learning About Assessment, Learning Through Assessment*. Online: www.nap.edu

Finch, F.L., editor. 1991. *Educational Performance Assessment*. Chicago: Riverside Publishing Company/Houghton Mifflin.

Fleischman, H.L., and L. Williams. 1996. *An Introduction to Program Evaluation for Classroom Teachers*. Arlington, VA: Development Associates, Inc. Online: www.teacherpathfinder.org/School/Assess/assess.html

Fox, R.F. 1993. "Do Our Assessments Pass the Test?" *The School Administrator* Vol. 50, No. 11: 6.

Frederiksen, J.R., and A. Collins. 1989. "A Systems Approach to Educational Testing." *Educational Researcher* Vol. 18, No. 9: 27-32.

Gardner, H. 1989. "Assessment in Context: The Alternative to Standardized Testing." In *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction*, edited by B. Gifford. Boston, MA: Kluwer Academic Press.

Gardner, H. 1983. *Frames of Mind*. New York: Basic Books.

Glickman, C. 1991. "Pretending Not to Know What We Know." *Educational Leadership* Vol. 48, No. 8: 4-10.

Hall, R., J. Knudsen, and J. Green. 1995/1996. "A Case Study of Systemic Aspects of Assessment Technologies." *Educational Assessment* Vol. 3, No. 4: 315-361.

Haney, W. 1985. "Making Testing More Educational." *Educational Leadership* Vol. 43, No. 2: 4-13.

Hansen, J.B., and W.E. Hathaway. 1993. *A Survey of More Authentic Assessment Practices*. Washington, DC: The ERIC Clearinghouse on Tests, Measurement, and Evaluation.

Hart, D. 1994. *Authentic Assessment: A Handbook for Educators*. Menlo Park, CA: Addison-Wesley Press.

Henning-Stout, M. 1994. *Responsive Assessment: A New Way To Think About Learning*. San Francisco: Jossey-Bass Publishers.

Herman, J., P. Aschbacher, and L. Winters. 1992. *A Practical Guide to Alternative Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Hibbard, K.M., and M.E. Yakimowski. 1997. *Assessment in Connecticut: A Partnership to Improve Student Performance—Connecting State-Level Assessment and Classroom Practices*. Cheshire, CT: Connecticut Association for Supervision and Curriculum Development.

Hopstock, P., M.B. Young, and A.M. Zehler. 1993. *Serving Different Masters: Title VII Evaluation Practice and Policy* (Volume 1, Final Report). Report to the U.S. Department of Education, Office of Policy and Planning. Arlington, VA: Development Associates, Inc.

Johnston, P., S. Guice, K. Baker, J. Malone, and N. Michelson, N. 1995. "Assessment of Teaching and Learning in Literature-based Classrooms." *Teaching and Teacher Education* Vol. 11, No. 4: 359-371.

Jones, R. 1998. "What Works: Researchers Tell What Schools Must Do to Improve Student Achievement." *American School Board Journal* (April 1998). Online: www.asbj.com/199804/0498coverstory.html

Jorgensen, M. 1993. *Assessing Habits of Mind: Performance-Based Assessment in Science and Mathematics*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.

Kean, M.H. Undated. Educational Assessment: Four Principles to Consider. Online: www.ctb.com about_assessment/four_principles.shtml

Killion, J., and G.T. Bellamy. 2000. "On the Job: Data Analysts Focus School Improvement Efforts." *Journal of Staff Development* (Winter 2000). Online: www.nsdc.org/educatorindex.htm

Kulm, G. 1994. *Mathematics Assessment: What Works in the Classroom.* 1st edition. San Francisco: Jossey-Bass Publishers.

Love, N. 2001. "The Heart of Systemic Reform." *ENC Focus* Volume 8, Number 2: 14-15.

Maeroff, G. 1991. "Assessing Alternative Assessment." *Phi Delta Kappan* Vol. 73, No. 4: 272-281.

Marzano, R., D. Pickering, and J. McTighe. 1993. *Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R., and J.S. Kendall. 1996. *A Comprehensive Guide to Designing Standards-Based Districts, Schools and Classrooms.* Alexandria, VA: Association for Supervision and Curriculum Development.

McTighe, J. 1996. "What Happens Between Assessments." *Educational Leadership* (December 1996/ January 1997). Online: www.ascd.org/readingroom/edlead/9612/mctighe.html

Mehrens, W. 1992. "Using Performance Assessment for Accountability Purposes." *Educational Measurement: Issues and Practices* Vol. 11, No. 1: 3-9, 20.

Mills, R. 1989. "Portfolios Capture Rich Array of Student Performance." *The School Administrator* Vol. 11, No. 46: 8-11.

Ministry of Education, Province of British Columbia. 1993. *Supporting Learning, Understanding and Assessing the Progress of Children in the Primary Program.* Toronto, BC: Author.

Minnesota Department of Children, Families, and Learning. Undated. *Continuous Improvement.* Online: cfl.state.mn.us/cip/process.html

Mitchell, R. 1992. *Testing for Learning.* New York: Free Press/MacMillan.

Morris, L., L. Fitz-Gibbon, and C.T. Lindheim. 1987. *How to Measure Performance and Use Tests.* 11th printing. (Series: Program Evaluation Kit, Volume 7.)

National Forum on Assessment. 1995. *Principals and Indicators for Student Assessment Systems.* Cambridge, MA: FairTest.

Neill, D.M. Undated. *Transforming Student Assessment.* Online: www.fairtest.org/MNKappan.html

Newmann, F., W. Secada, and G. Wehlage. (1998) *A Guide to Authentic Instruction and Assessment: Vision, Standards and Scoring.* Madison, WI: Wisconsin Center for Education Research.

North Carolina Department of Education, Department of Public Instruction. 1999. *Parents' Guide to Classroom Assessment.* Online: www.ncpublicschools.org/state_board/

North Carolina Department of Public Instruction, Evaluation Section, Division of Accountability Services. 2000. *Closing the Achievement Gap: Views from Nine Schools.* Online: www.ncpublicschools.org/closingthegap/nineschools.pdf.

Northwest Regional Educational Laboratory. 1998a. "Assessment—Where the Rubber Meets the Road in Standards-Based Education." In *Toolkit 98.* Online: www.nwrel.org/assessment/toolkit98/rubber.html

Northwest Regional Educational Laboratory. 1998b. "Grading and Reporting: A Closer Look." In *Toolkit 98.* Online: www.nwrel.org/assessment/toolkit98/chapter4.html

Nickse, R., et al. 1981. *Competency-Based Education.* New York: Teachers College Press.

O'Neil, J. 1990. "Piecing Together the Restructuring Puzzle." *Educational Leadership* Vol. 47, No. 7: 4-10.

Ormrod, J.E. 2000. *Educational Psychology: Developing Learners, 3rd edition.* Columbus, OH: Merrill.

Perrone, V., editor. 1991. *Expanding Student Assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Peterson, Kent. 1995. "Critical Issue: Creating High-Achieving Learning Environments." North Central Regional Educational Laboratory. Online: www.ncrel.org/sdrs/areas/issues/educatrs/leadrshp/le400.htm

Philadelphia Education Fund. Undated. *Guidelines for Looking at Student Work.* Online: www.philaedfund.org/slcweb/guideli.htm

Pimentel, S., and L.A. Arsht. 1998. "Don't Be Confused by Rankings; Focus on Results." *Education Week* (November 11, 1998). Online: www.edweek.org/ew/vol-18/11arsht.h18

Popham, W.J. 1995. *Classroom Assessment: What Teachers Need to Know.* Boston, MA: Allyn and Bacon.

Rabinowitz, S.N., and S. Ananda. 2001. *Balancing Local Assessment with Statewide Testing: Building a Program that Meets Student Needs.* San Francisco, CA: WestEd. Online: web.wested.org/online_pubs/kn-01-01.pdf

Rothman, R. 1995. *Measuring Up: Standards, Assessment, and School Reform.* San Francisco: Jossey-Bass, Publishers.

Rothman, R. 2000. *Bringing All Students to High Standards: Report on National Education Goals Panel Field Hearings.* Washington, DC: National Education Goals Panel.

Saskatchewan School Trustees Association Research Centre. 1992. Drawing Value from Evaluation. Regina, Saskatchewan, Canada. Online: www.ssta.sk.ca/research/evaluation_and_reporting/9d

Sanders, W.L. 1998. "Value-Added Assessment." *AASA School Administrator* (December 1998). Online: www.aasa.org/publications/sa/1998_12/sanders.htm

Schafer, W.D., et al. Undated. *Study of Higher-Success and Lower-Success Elementary Schools.* Online: www.mdk12.org/practices

Scriven, M. 1967. "The Methodology of Evaluation." In *Curriculum Evaluation,* R.E. Stake, editor (American Educational Research Association Monograph Series on Evaluation, No. 1, pp. 39-83). Chicago: Rand McNally.

Shanker, A. 1989. "A Good Job for Good Grades." *The New York Times* (March 5, 1989).

Stiggins, R.J. 1997. *Student-centered Classroom Assessment.* 2nd edition. Upper Saddle River, NJ: Merrill.

Stiggins, R.J., and N. Conklin. 1992. *In Teachers' Hands: Investigating the Practices of Classroom Assessment.* Albany, NY: State University of New York Press.

U.S. Department of Education. 1996. "What the Research Says about Student Assessment." *Improving America's Schools* (Spring 1996). Online: www.ed.gov/pubs/IASA/newsletters/assess/pt4.html

WestEd and National Council of Teachers of Mathematics. Undated. *Learning from Assessment—Tools for Examining Assessment Through Standards: A Professional Development Resource for Middle School Mathematics Teachers.* Authors. Online: www.wested.org/lfa/about/professional_development.htm

Wiggins, G. 1993. *Assessing Student Performance: Exploring the Purpose and Limits of Assessment.* Jossey-Bass Education Series, Vol. 1. San Francisco, CA: Jossey-Bass Publishers.

Wiggins, G. 1991. "Standards, Not Standardization: Evoking Quality Student Work." *Educational Leadership* Vol. 48, No. 5: 18-25.

Wiggins, G. 1989. "A True Test: Toward More Authentic and Equitable Assessment." *Phi Delta Kappan* Vol. 70, No. 9: 703-713.

Wolf, D., J. Bixby, J. Glenn, and H. Gardner, editors. 1991. *To Use Their Minds Well: Investigating New Forms of Student Assessment.* Washington, DC: American Educational Research Association.

National Principals Resource Center
1615 Duke Street
Alexandria, VA 22314
703-684-3345
www.naesp.org/nprc

Educational Research Service
2000 Clarendon Boulevard
Arlington, VA 22201
703-243-2100

BEST COPY AVAILABLE